

**MESTRADO**

MULTIMÉDIA - ESPECIALIZAÇÃO EM TECNOLOGIAS

# ***Design* de Usabilidade em Interfaces Conversacionais Híbridas**

Ana Sofia Ferreira de Sousa

**M**

2017

FACULDADES PARTICIPANTES:

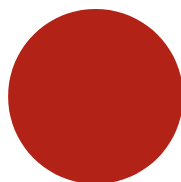
FACULDADE DE ENGENHARIA

FACULDADE DE BELAS ARTES

FACULDADE DE CIÊNCIAS

FACULDADE DE ECONOMIA

FACULDADE DE LETRAS







# ***Design* de Usabilidade em Interfaces Conversacionais Híbridas**

**Ana Sofia Ferreira de Sousa**

Mestrado em Multimédia da Universidade do Porto

Orientadora: Professora Doutora Maria Teresa Magalhães da Silva Pinto de Andrade

17 Julho de 2017



©Ana Sofia Ferreira de Sousa, 2017

***Design* de Usabilidade em Interfaces**  
**Conversacionais Híbridas**

**Ana Sofia Ferreira de Sousa**

Mestrado em Multimédia da Universidade do Porto

Aprovado em provas públicas pelo Júri:

Presidente: Professor Doutor André Monteiro de Oliveira Restivo

Vogal Externo: Professor Doutor António Joaquim da SilvaTeixeira

Orientadora: Professora Doutora Maria Teresa Magalhães da Silva Pinto de Andrade



# Resumo

Esta dissertação retrata essencialmente o *design* de usabilidade em Interfaces Conversacionais que permitem ao utilizador interagir através de linguagem natural. Estas interfaces apesar de evoluírem de forma constante e apresentarem uma maior tendência em relação à adesão por parte de empresas de tecnologia, expõem ainda diversos problemas e limitações no que diz respeito à usabilidade.

Assim, o desafio da presente dissertação envolve a aplicação de *e-commerce* Shelf AI, desenvolvida pela Xarevision, que permite ao utilizador realizar compras, procurando o produto através de texto, voz ou código de barras. Apesar de ser possível a pesquisa por voz, esta não inclui a capacidade de o utilizador interagir através de um diálogo natural. Posto isto, o objetivo passa por transformar a Shelf AI numa interface conversacional híbrida, onde o utilizador tenha a possibilidade de interagir de forma híbrida, ou seja, por voz e por toque/texto.

No caso de interfaces conversacionais, a usabilidade depende maioritariamente do fluxo de diálogo, ou seja, de como o utilizador interage com o sistema através de linguagem natural. O processo de *design* foca-se principalmente no fluxo de diálogo, onde é necessário compreender o utilizador, existir uma compreensão clara dos fatores humanos assim como outros fatores que possam interferir negativamente com a usabilidade do sistema. Tendo em conta que a Shelf AI está a ser convertida numa interface híbrida, foram delineadas novas funcionalidades que conduziram à alteração ou novas modificações em relação à componente gráfica assim como a conversacional.

Palavras-chave: *Design* de Usabilidade; Interfaces Conversacionais; Interfaces Conversacionais Híbridas; Interação Humano-Computador; Diálogo Natural; *E-commerce*





# Abstract

This dissertation portrays essentially the usability design in Conversational Interfaces that allows the user to interact through natural language. Even though these interfaces evolve steadily and present a greater tendency in relation to accession by technology companies, there are still several problems and limitations when it comes to usability.

Thus, the challenge of this dissertation involves the application of e-commerce Shelf AI, developed by Xarevision, which allows the user to make purchases, look for the product via text, voice or barcode. Although it is possible to search by voice, this does not include the ability for the user to interact through a natural dialogue. Having said that, the goal is to turn the Shelf AI into a hybrid conversational interface, where the user has the possibility to interact in hybrid form, i.e. by touch and voice/text.

In the case of conversational interfaces, usability depends on mostly dialog flow, which is how the user interacts with the system through natural language. The design process focuses mainly on the dialog flow, where it is necessary to understand the user, a clear understanding of human factors as well as other factors that may interfere negatively with the usability of the system. Taking into account that the Shelf AI is being converted into a hybrid interface, new features were defined which led to the amendment or new modifications in relation to the graphic component as well as conversational.

**Keywords:** Usability Design; Conversational Interfaces; Hybrid Conversational Interfaces; Human-Computer Interaction; Natural Dialogue; E-commerce



# Agradecimentos

A concretização desta dissertação só foi possível com a ajuda e apoio de algumas pessoas das quais preciso dedicar palavras de gratidão especialmente a elas.

À minha família e especialmente ao meus pais pelo apoio e todos os sacrifícios que passaram para poder atingir os meus objetivos no ensino superior.

À minha irmã um especial obrigado por todo o apoio dado, pelo exemplo que tenho a seguir, pelo pilar que foi nos momentos mais difíceis e por toda a paciência.

À orientadora Professora Doutora Maria Teresa Andrade pelo tempo disponibilizado, pela sua orientação e ajuda.

À equipa da Xarevision por toda a disponibilidade, ajuda e incentivo demonstrados e especialmente ao Engenheiro Sílvio Macedo e Norberto Amaral pela total disponibilidade em ajudar, pelo grande incentivo e pela supervisão e acompanhamento semanal pois foram imprescindíveis para atingir os prazos definidos.

Ao meu namorado por toda a paciência, compreensão e incentivo dado para nunca recuar durante este longo período.

À minha amiga e colega Sara Oliveira por todas as horas que partilhamos de trabalho e pelo apoio demonstrado, que foi muito importante nos momentos mais difíceis de todo o trabalho.

A todos os outros amigos por todo o apoio demonstrado durante este período.



# Índice

1. Introdução .....	1
1.1. Contexto e Motivação .....	1
1.2. Apresentação da Organização .....	2
1.3. Projeto: Shelf AI .....	2
1.4. Problemas e Objetivos .....	3
1.5. Metodologia de Investigação .....	4
1.6. Estrutura da Dissertação .....	5
2. Estado de Arte.....	7
2.1. Interfaces Conversacionais .....	7
2.1.1. A Evolução das Interfaces Conversacionais .....	8
2.1.2. <i>Chatbots</i> .....	10
2.1.3. Assistentes de Voz .....	13
2.1.4. Interfaces Human-powered .....	15
2.1.5. Comércio Conversacional .....	17
2.2. Design de Usabilidade .....	19
2.2.1. Interação Humano-Computador .....	19
2.2.2. Usabilidade .....	20
2.2.3. Protótipos .....	22
2.3. Design em Interfaces Conversacionais .....	23
2.4. Análise de Aplicações Relacionadas .....	29
2.4.1. Google Express .....	30
2.4.2. Instacart.....	31
2.4.3. FreshDirect.....	33
2.4.4. Shipt.....	35
2.4.5. Síntese da Análise .....	38
2.5. Síntese do Capítulo.....	38
3. Análise e Design.....	40
3.1. Questionário Inicial .....	40

3.2. Público-alvo .....	47
3.3. Fatores Diferenciadores .....	48
3.4. Funcionalidades .....	49
3.5. Fluxos de Interação .....	51
3.6. Design de Usabilidade Conversacional .....	54
3.6.1. Alterações na Aplicação .....	55
3.6.2. Fluxos de diálogo .....	59
3.7. Síntese do Capítulo.....	71
4. Definição dos Protótipos e Testes .....	74
4.1. Desafios nas Tecnologias .....	74
4.2. Tecnologias e Métodos Utilizados .....	77
4.3. Testes de usabilidade.....	79
4.3.1. Objetivos dos Testes.....	79
4.3.2. Perfil dos participantes .....	80
4.3.3. Tarefas.....	80
4.3.4. Resultados .....	82
4.3.5. Questionário Pós-Teste.....	83
4.4. Síntese do Capítulo.....	90
5. Conclusões e Trabalho Futuro.....	91
6. Bibliografia .....	94
7. Anexos.....	100
Anexo A – Guião do Teste de Usabilidade.....	100
Anexo B – Inquérito Pós-Teste de Usabilidade .....	102
Anexo C – Outros Fluxos de Diálogo .....	104

# Lista de Figuras

Figura 1. Resultados do produto pesquisado de acordo com a semântica da Shelf AI.	3
Figura 2. Um dos primeiros <i>posts</i> inofensivos de Tay.	11
Figura 3. Exemplo de um post ofensivo de Tay.	11
Figura 4. Zo tem capacidade de evitar certos temas como a Política.	12
Figura 5. Iniciativa própria de Zo para interagir com o utilizador.	12
Figura 6. Processo de compra na plataforma Magic.	16
Figura 7. Simulação de proximidade entre o agente e o utilizador.	17
Figura 8. Utilização crescente das aplicações de mensagens (BI Intelligence , 2016).	18
Figura 9. Primeiro protocolo de interação	27
Figura 10. Segundo protocolo de interação	27
Figura 11. Terceiro protocolo de interação	28
Figura 12. Página <i>Home</i> da aplicação Google Express.	30
Figura 13. Filtragem de produtos da aplicação Google Express.	31
Figura 14. Página <i>Home</i> da aplicação Instacart.	32
Figura 15. Carrinho de compras da aplicação Instacart.	33
Figura 16. Página <i>Home</i> da plataforma FreshDirect	34
Figura 17. Carrinho de compras da plataforma FreshDirect.	35
Figura 18. Página <i>Home</i> da aplicação Shipt.	36
Figura 19. Resultados surgidos durante a entrada do utilizador	37
Figura 20. Representação das faixas etárias dos participantes do questionário	41
Figura 21. Número de utilizações de interfaces conversacionais	42
Figura 22. Preferência do utilizador de como interagir com a interface	43
Figura 23. Preferência dos utilizadores do modo de interagir do sistema	44
Figura 24. Nível de interesse dos utilizadores quanto a compras através de um diálogo	45
Figura 25. Nível de utilidade de uma interface conversacional quanto a compras	46
Figura 26. Início de sessões na Shelf AI	52
Figura 27. Ativar ou desativar de interação por voz	52
Figura 28. Pesquisa de produtos independentemente através de voz ou texto	53
Figura 29. Processo de <i>checkout</i>	54
Figura 30. <i>Feedback</i> da Shelf AI original no reconhecimento de voz	56

Figura 31. Alternativa ao <i>feedback</i> através de <i>pop ups</i>	57
Figura 32. <i>Pop up</i> do <i>checkout</i> original da Shelf AI	58
Figura 33. Ecrã do processo de <i>checkout</i>	59
Figura 34. Exemplo de fluxo de diálogo para a pesquisa do produto vinho	62
Figura 35. Exemplo de um momento de pesquisa filtrada e com resultados específicos	63
Figura 36. Fluxo de diálogo de boas vindas	64
Figura 37. Sugestão do sistema para uma pesquisa de produtos mais rápida	64
Figura 38. Exemplo de uma pesquisa completa em relação ao produto	65
Figura 39. Interação entre o utilizador e o sistema no carrinho de compras	66
Figura 40. Processo de <i>checkout</i> com diferentes tipos de interação e escolhas	68
Figura 41. Dúvida por parte do utilizador com respostas associadas	69
Figura 42. Fluxo de pesquisa que permite detetar informação importante a armazenar	71
Figura 43. Nível de naturalidade da fala do sistema	85
Figura 44. Nível de utilidade dos tutoriais para iniciantes	85
Figura 45. Número de respostas quanto ao grau de satisfação dos utilizadores	86
Figura 46. Nível de vantagem de uma interface híbrida em comparação a uma gráfica	86
Figura 47. Grau de utilidade da interação por voz para diferentes contextos	87



# Lista de Tabelas

Tabela 1. Comparação de acordo com alguns critérios das aplicações analisadas.	38
Tabela 2. Questões de escolha múltipla e quantitativas	42
Tabela 3. Vantagens associadas às interfaces híbridas para compras	47
Tabela 4. Funcionalidades do ponto de vista do utilizador	50
Tabela 5. Funcionalidades do ponto de vista do sistema	51
Tabela 6. Filtros definidos para os primeiros produtos	60
Tabela 7. <i>Prompts</i> iniciais e gerais de acordo com o tipo de filtro	61
Tabela 8. Ferramentas para simular e/ou implementar conversações	75
Tabela 9. Tarefas do teste de usabilidade	81
Tabela 10. Questões quantitativas da primeira secção	84
Tabela 11. Questões qualitativas da segunda secção	87
Tabela 12. Respostas do que mais agradou ao utilizador na Shelf AI	88
Tabela 13. Sugetões de melhorias dos utilizadores em relação à Shelf AI híbrida	89



# Abreviaturas e Símbolos

A.L.I.C.E	Artificial Linguistic Internet Computer Entity
AIML	Artificial Intelligence Markup Language
HCI	Human-Computer Interaction
MIT	Massachusetts Institute of Technology
UML	Unified Modeling Language
UX	User Experience
STT	Speech to Text
TTS	<i>Text to Speech</i>
WoZ	<i>Wizard of Oz</i>



# 1. Introdução

No presente capítulo de carácter introdutório, será abordado o âmbito e o contexto desta dissertação, a análise do problema a resolver, assim como os objetivos que são pretendidos atingir com a realização da dissertação. Posteriormente, também é apresentada a metodologia de investigação adotada e uma breve síntese de toda a estrutura da dissertação.

## 1.1. Contexto e Motivação

Atualmente os humanos recorrem aos dispositivos tecnológicos para poderem aceder a qualquer informação e realizar as mais variadas tarefas. No entanto, ao longo dos anos diversas áreas na tecnologia têm evoluído, alterando assim aos poucos os meios de interação com as máquinas. O objetivo de Alan Turing de provar que as máquinas podem pensar e simular o comportamento humano, levou ao aparecimento de outras interfaces conversacionais na competição Loebner Prize<sup>1</sup>. Porém, a existência delas só é possível devido à área de Inteligência Artificial e Processamento de Linguagem Natural que estão em constante evolução.

Estas interfaces apresentam um grande desafio na área de *design* comparativamente ao *design* de interfaces gráficas. Esse desafio passa por compreender como interagir com o utilizador, não podendo recorrer a conteúdo visual para conduzir a interação. Por conseguinte, é necessário compreender várias áreas de forma a otimizar a interação neste tipo de interfaces.

As empresas começaram a criar interfaces conversacionais de forma a desenvolverem novos serviços e técnicas para estarem mais disponíveis assim como para aumentar o número de vendas.

Apesar de estas interfaces estarem em constante evolução na área de Inteligência Artificial e Processamento de Linguagem Natural, em alguns casos a usabilidade do sistema ainda não está bem presente. A usabilidade do sistema muitas vezes não depende da tecnologia, mas sim do utilizador, por isso existe a necessidade de estudá-lo aprofundadamente. Assim, é necessário

---

<sup>1</sup> Concurso na área de inteligência artificial para vencer o *chatbot* considerado mais humano.

existir uma compreensão de fatores humanos como a memória de curto prazo e audição, assim como os limites e capacidades das tecnologias.

## **1.2. Apresentação da Organização**

A Xarevision é uma empresa criada em 2006 e que se dedica essencialmente às redes digitais de *displays* interativos e estáticos, sendo líder em tecnologias para o retalho. Esta liderança é assumida pelo facto dos seus projetos atingirem mais de 40% da população ativa em Portugal.

O seu primeiro projeto foi desenvolvido para o canal UPMEDIA, da Universidade do Porto, e pertencia à área de TV corporativa onde a empresa tinha como objetivo gerir, distribuir e entregar os conteúdos em diferentes locais pela Universidade. Desde então, a Xarevision centra-se em comunicar com tecnologia cada vez mais avançada, recorrendo assim às redes sem fio e inteligência artificial.

Atualmente a Xarevision está a apostar nas interfaces conversacionais, pois devido aos avanços tecnológicos nas áreas de inteligência artificial e processamento de linguagem natural, estas interfaces são cada vez mais comuns no *e-commerce*. Por conseguinte, surgiu o projeto da Shelf AI que se encontra em fase de desenvolvimento.

A empresa desenvolve ainda produtos para gerir filas, sendo que estes são testados para analisar e aumentar a qualidade do atendimento ao cliente. Além disto, desenvolve redes digitais de *displays* para transmitir ao público a informação desejada pelo cliente. A Xarevision dedica-se também à criação de media interativa como quiosques, com informação de carácter interativo e de ecrã tátil.

## **1.3. Projeto: Shelf AI**

Esta dissertação terá como base o projeto a Shelf AI, uma plataforma de *e-commerce* baseada em Inteligência Artificial e Interface Conversacional. Com a Shelf AI, as empresas de retalho alimentar podem melhorar a experiência de compra dos seus clientes, com vista ao aumento das vendas através das suas aplicações móveis. O objetivo é permitir uma interação mais rápida e fácil entre clientes e retalhistas.

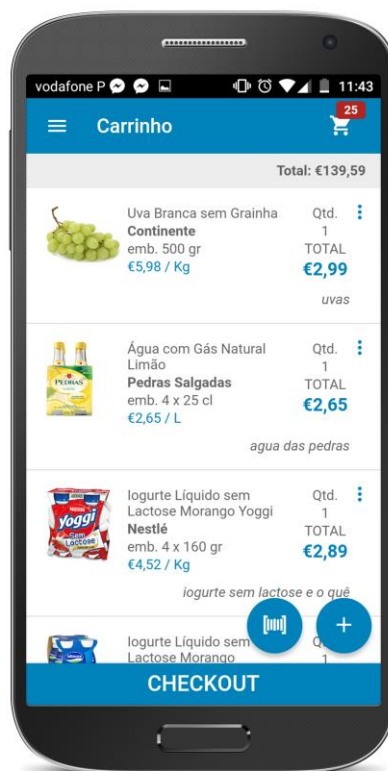


Figura 1. Resultados do produto pesquisado de acordo com a semântica da Shelf AI.

A Shelf AI estará inserida nas aplicações de supermercados, não sendo necessário os clientes instalarem outra aplicação. A plataforma para além de estar disponível nos *smartphones*, também pode ser usada em interfaces como das Assistentes de Voz Alexa, armazenada no dispositivo Amazon Echo, e Cortana.

A plataforma recorre a Inteligência Artificial para personalizar a experiência dos utilizadores e tem a capacidade de aprender as suas preferências em relação a produtos e marcas, apresentando resultados optimizados para o utilizador. É composta por uma interface conversacional que se baseia no reconhecimento da voz natural do cliente, bastando para o seu *smartphone* para realizar as suas compras. Esta plataforma facilitará o processo de compra, diminuindo a necessidade de o utilizador recorrer ao toque para interagir com a *app* móvel.

## 1.4. Problemas e Objetivos

A presente dissertação tem como objetivo o estudo das diversas interfaces conversacionais, assim como a sua constante evolução, não só em relação às funcionalidades mas também em relação à usabilidade com que o utilizador se depara durante a interação.

Estas interfaces sendo baseadas em linguagem natural, torna-se necessário fazer um estudo dos fatores humanos que influenciam a experiência, assim como as necessidades dos utilizadores. É importante entender como ocorre a interação entre humanos, essencialmente em contexto de compras, e como estes interagem com os sistemas que providenciam a oportunidade de diálogo e interação. Outro estudo passa por adquirir o conhecimento das regras e princípios do *design* a ter em conta para potenciar a uma melhor usabilidade deste tipo de interfaces. Este mesmo estudo focar-se-á em interfaces conversacionais que funcionam maioritariamente por voz e não por texto (como os *chatbots*). Desta forma vai ser possível a obtenção dos conhecimentos necessários para o projeto Shelf AI, que se foca na interação por voz do utilizador com a plataforma para a realização de compras.

Atualmente, a aplicação Shelf AI funciona com pesquisa através de voz e texto, enquanto os resultados da interface são apresentados visualmente, isto é, somente por texto. Um dos objetivos passa por transformar esta interface de modo a que seja possível a pesquisa por voz e texto e que comunique os resultados da mesma forma, ou seja, transformá-la numa interface híbrida. A finalidade será avaliar a eficácia, a importância do *design* de usabilidade e a utilidade de uma interface conversacional híbrida numa plataforma direcionada a compras (Shelf AI). Deste modo, serão criados protótipos da Shelf AI e inquiridos após o teste para compreender e concluir se uma interface híbrida é a melhor opção para o utilizador experienciar uma interação completa e intuitiva e se vai ao encontro às suas necessidades.

Em suma, a dissertação é realizada tendo como base as seguintes questões de investigação:

1. Como é que as interfaces conversacionais podem ser úteis para o utilizador?
2. Como é a interação entre humanos em contexto de compras?
3. Quais as expectativas de resposta por parte do utilizador das interfaces conversacionais *e-commerce*?
4. Como é que a usabilidade pode otimizar a experiência e interação do utilizador?
5. Quais os objetivos e necessidades do utilizador ao recorrer a uma interface conversacional *e-commerce*?

## 1.5. Metodologia de Investigação

A dissertação começa por uma investigação e escrita do estado de arte relacionado com a área de Interfaces Conversacionais e o *Design* de Usabilidade. De seguida, para a concepção do projeto, de um certo modo tentando contribuir para os problemas referidos na dissertação, surgem então outras etapas.

A primeira etapa consiste em realizar uma investigação que aborde as áreas de HCI (Human-Computer Interaction), usabilidade e UX (User Experience) *Design*. Esta etapa também terá em



conta as análises já realizadas no estado de arte de projetos semelhantes. Deste modo, serão realizados inquéritos a potenciais utilizadores para compreender as necessidades e quais as opiniões que têm em relação a interfaces conversacionais.

A segunda etapa será para colocar em prática o processo de *design* de interação através de três fases. Na primeira fase de requisitos, será definido o que é necessário para a interface, recorrendo às análises já realizadas na primeira etapa, assim como a inquéritos aos potenciais utilizadores. Também será analisada toda a informação relevante para determinar como funcionará a Shelf AI em relação ao diálogo com o utilizador. Deste modo, a segunda fase consistirá na análise da informação recolhida com o intuito de detetar problemas das interfaces já existentes. Por fim, será possível passar à fase de *design*, aliando as necessidades dos utilizadores aos princípios de *design*. Serão assim criados protótipos para serem testados para detetar eventuais problemas com a interface e então resolvê-los com a criação de outro protótipo. Este processo de repetição do ciclo denomina-se por *design* iterativo, e o seu objetivo é repetir as vezes necessárias o processo mas acontecer apenas nos problemas que valem a pena de serem corrigidos.

## 1.6. Estrutura da Dissertação

Esta dissertação divide-se em cinco capítulos. No presente e primeiro capítulo da Introdução é abordado o tema e o contexto em que a dissertação se insere e são definidos os objetivos e questões de investigação que servem como base.

No capítulo 2 do Estado de Arte é realizada uma revisão à literatura em relação às interfaces conversacionais, uma breve análise às que mais se destacam atualmente e uma revisão dos aspetos relacionados com o *design* e usabilidade relacionados principalmente com as interfaces conversacionais. Também são analisadas várias interfaces que se enquadram com o tema de *e-commerce* para tomar conhecimento de como estas funcionam em relação à usabilidade e quais as suas funcionalidades.

No capítulo 3 da Análise e Design é apresentado um estudo inicial com base num questionário realizado aos potenciais utilizadores, onde é traçado o perfil do público alvo e apresentado o conhecimento e opinião dele sobre as interfaces conversacionais. Posteriormente, são expostos os fatores diferenciadores e funcionalidades da Shelf AI híbrida. Por fim, neste capítulo são abordadas todas as

são apresentados os inquéritos realizados para os potenciais utilizadores assim como os seus resultados. Com isto, é possível passar para a definição das funcionalidades e requisitos da Shelf AI.

No capítulo 4 serão escolhidas as metodologias de *design*, avaliação de usabilidade e o tipo de protótipo para serem utilizados nos testes de usabilidade abordados neste capítulo.

No capítulo 5 é realizada uma análise da Shelf AI de acordo com os resultados obtidos de forma chegar a alguma conclusão.

Por fim, no último capítulo, serão apresentas as conclusões da realização desta dissertação e as possíveis perspectivas para o trabalho futuro.

## 2. Estado de Arte

Neste capítulo são essencialmente apresentados trabalhos relacionados com interfaces conversacionais, é abordada um pouco da história e evolução destas interfaces assim como os diferentes tipos destas interfaces e aspetos relacionados com o *design* de usabilidade.

### 2.1. Interfaces Conversacionais

O avanço da tecnologia tem permitido o desenvolvimento de sistemas conversacionais que permitem o uso de linguagem natural para interagir (Rudnicky & Hauptmann, 1989). As interfaces conversacionais são sistemas em que o utilizador pode interagir por voz e/ou texto através de um diálogo natural. De acordo com Kaplan (2013), estas interfaces são inteligentes porque além de funcionarem com fala sintetizada e reconhecimento de fala, conseguem processar a informação separadamente e ainda compreender o significado da entrada do utilizador. Ainda assim, têm também a capacidade de interagir de acordo com o contexto, dependendo do utilizador e do momento em que é necessário (Kaplan, 2013).

Estas interfaces são sistemas que recebem informação em linguagem natural, proveniente do utilizador, convertendo-a de forma a estruturar, processar, agir e criar respostas naturais (Shapiro, 2016),

Segundo (Allen, D. Byron, Ferguson, Galescu, & Stent, 2000) os investigadores definem as interfaces conversacionais de forma diferente, mas todos defendem que para uma interface deste tipo é necessário existir a interação com um humano. Evidenciam que para alguns, uma interface destas resume-se a uma estrutura de interação na qual são especificadas as mensagens do sistema e através destas, as repostas do utilizador são restringidas. Para outros investigadores, uma interface conversacional tem a capacidade de imitar um diálogo de humano para humano.

Todos os sistemas de diálogo podem ser denominados por sistemas conversacionais mas são distinguidos de acordo com intensidade com que o sistema controla ou não o diálogo e a interação do utilizador para com este (Glass, Weinstein, Cyphers, & Polifroni, 2005). Os diálogos que acontecem nestes sistemas podem ter diferentes tipos de iniciativas, dependendo do responsável

que direciona o diálogo e são estes: iniciativa do utilizador, iniciativa do sistema e iniciativa mista (Meng, Ching, Chan, Wong, & Chan, 2004).

Por iniciativa do utilizador entende-se que o diálogo é controlado unicamente pelo utilizador e o sistema mantém um papel passivo conseguindo satisfazer as expectativas deste durante a interação. Isto pode levar à frustração do utilizador, caso o sistema não esteja preparado para responder a certos temas e vocabulários. O diálogo por iniciativa do sistema acontece quando o diálogo é somente controlado pelo sistema, ou seja, este é capaz de dominar o fluxo do diálogo, restringir a interação do utilizador e definir quando o utilizador pode realizar uma entrada. A iniciativa mista proporciona uma interação dinâmica onde tanto o utilizador como o sistema têm controlo sob o diálogo (Meng, Ching, Chan, Wong, & Chan, 2004)

Uma interface conversacional é composta por *prompts*, gramática e fluxo da interação. As *prompts* são mensagens do sistema que funcionam como instruções durante a interação. A gramática é responsável por definir e limitar o que o sistema pode compreender nas entradas do utilizador. O fluxo determina quais as ações que o sistema pode tomar de acordo com as ações do utilizador (Cohen, Giangola, & Balogh, 2004).

Existem dois tipos de interfaces conversacionais distintos: os *Chatbots*, onde a interação é predominantemente por texto, e os Assistentes de Voz que, como o próprio nome indica, a voz é o principal meio de interação.

### **2.1.1. A Evolução das Interfaces Conversacionais**

Alan Turing foi um matemático considerado um dos criadores da Ciência dos Computadores, enfrentando obstáculos para ultrapassar as fronteiras da inteligência da máquina.

A máquina de Turing foi criada antes do aparecimento do computador e consiste num modelo matemático que está relacionado com os limites da computabilidade. Esta máquina pode ser vista como um estado de máquina, pois consegue estar num determinado estado e transitar para outro através de um processo com instruções para a máquina (Barker-Plummer, 2016).

Alan Turing em 1950 apresentou o Jogo da Imitação, um teste para responder à questão "*Can machines think?*". Este teste tinha como objetivo que a máquina simulasse a capacidade de pensar como um humano. A ideia inicial passava por definir os termos "*machine*" e "*think*" mas devido à sua ambiguidade, Turing procurou solucionar este problema recorrendo a um jogo – o Jogo da Imitação.

O jogo consistia em três participantes: um Homem, uma Mulher e um interrogador. O interrogador encontrava-se numa sala diferente dos outros participantes, para delinear as capacidades físicas e intelectuais. O objetivo passava por testar a capacidade de distinguir o Homem da Mulher, sem qualquer contacto direto com ambos. O interrogador tinha a oportunidade de fazer questões aos participantes para que as respostas pudessem ajudá-lo a

identificar cada participante. As perguntas e respostas eram feitas através da escrita, uma vez que através da voz a identificação tornava-se facilitada (Turing, 1950).

O primeiro sistema conversacional bem-sucedido, após o teste de Turing, foi ELIZA, desenvolvido por Joseph Weizenbaum (1966) e vencedor das três primeiras edições do Loebner Prize.

O programa não é composto por inteligência artificial pois limita-se a funcionar através de um grupo de regras de produção e manipulação do texto, dando a sensação de que é inteligente e que compreende o utilizador. Para o programa ser capaz de gerar as respostas idealizadas, tem de passar por uma série de problemas técnicos: identificar a palavra-chave – a palavra de maior importância, proveniente da entrada do utilizador; identificar o contexto de acordo com a palavra-chave escolhida anteriormente; produzir respostas através de mecanismos quando não existem palavras-chave; conter capacidade de editar e estender os scripts de ELIZA.

Devido a esta forma de gerar respostas, ELIZA não fica dependente de nenhuma língua, podendo através de um *script* ser implementada em diferentes línguas (Weizenbaum, 1966).

O script original de ELIZA passa por ser um *chatbot* que imita um psicoterapeuta Rogeriano<sup>2</sup>, mantendo um papel passivo durante o diálogo e dando a oportunidade ao utilizador de dirigir o tema e de se envolver na conversa. Durante o diálogo a ELIZA formula perguntas consoante as respostas do utilizador e contém várias frases definidas como “*Can you elaborate on that?*”, para conseguir manter o diálogo (Güzeldere & Franchi, 1995).

Na década seguinte surgiram mais tentativas para desenvolver sistemas conversacionais como é o caso do PARRY criado por Kenneth Colby em 1972. Este *chatbot* simula um esquizofrénico paranóico e é o oposto de ELIZA. Parry fala sobre os seus sentimentos, envolvendo o utilizador durante o diálogo em forma de pergunta-resposta (Güzeldere & Franchi, 1995). Além disto, Parry interpreta as respostas do utilizador, identificando um estado emocional e gera diversas respostas consoante as respostas anteriores e respeitando regras já especificadas (Colby, Hilf, Weber, & Kraemer, 1972).

Em meados de 1980, grande parte dos investigadores de ciência cognitiva e inteligência artificial não viam no futuro a possibilidade de criar um sistema com boas capacidades conversacionais. Existiam certos problemas que tinham de ser ultrapassados, como a dificuldade de processamento da linguagem natural. No entanto, os avanços tecnológicos desde a inteligência artificial até à ciência cognitiva, induziram investigadores a reexaminar a construção dos sistemas conversacionais (Graesser, et al., 2004). Segundo Cooper (2004), a visão de Turing em relação à computabilidade é simplificada e ainda utilizada.

A.L.I.C.E, surgida em 1995, é outro *chatbot* que se baseia no Jogo da Imitação de Alan Turing (1950). Este *chatbot* utiliza como linguagem a AIML própria para o estímulo-resposta necessário para os *chatbots*.

Em 2000 e 2001, A.L.I.C.E ganhou o Prémio Loebner pois os júris consideraram-na o computador “mais humano” apesar de as suas identificações, assim como os outros *chatbots* a concurso, não passarem para além de classificar humanos.

Apesar da semelhança da arquitetura de estímulo-resposta de de A.L.I.C.E com ELIZA, A.L.I.C.E distingue-se particularmente pois apresenta mais de 40.000 categorias de conhecimento em comparação que ELIZA apresentava apenas à volta de 200 categorias. Estas categorias são compostas por combinações de estímulos (perguntas) e respostas e a AIML tem a capacidade de armazenar os estímulos e corresponder determinadas respostas a este, através de uma estrutura em árvore (Wallace, 2009).

### **2.1.2. Chatbots**

Os *chatbots* são programas de computador que têm a capacidade de interagir com humanos através de linguagem natural. Compreendendo este tipo de linguagem, o programa recorre ao seu conhecimento para analisar assim as entradas dos utilizadores, de forma a gerar uma resposta também em linguagem natural. Os *chatbots* estão aptos para simularem um humano durante o diálogo, em forma de texto, e responderem às entradas dos utilizadores, podendo aproximar-se bastante de um humano (Shawar & Atwell, 2007). Estes permitem ao utilizador interagir e obter respostas através de texto, em qualquer momento, sem que seja necessário o apoio de um agente humano. Porém, uma das desvantagens passa por o utilizador não estar com as mãos livres para outras potenciais tarefas, que necessite de realizar.

A maioria dos *chatbots* atuais são criados e estão disponíveis em plataformas de *chat* famosas: Facebook Messenger, utilizado mundialmente, Kik usado sobretudo nos Estados Unidos e o WeChat com utilização dominante na China.

Posteriormente, nos subcapítulos seguintes (2.1.2.1 e 2.1.2.2) serão abordados dois exemplos de *chatbots* criados pela Microsoft, com o objetivo de estes adquirirem capacidades humanas através da aprendizagem por inteligência artificial.

#### **2.1.2.1. Tay**

Em março de 2016 a Microsoft lançou o *chatbot* Tay, com a finalidade ter conversas com os utilizadores simulando uma jovem humana através do Twitter. Tay foi programada para aprender e evoluir as suas capacidades de conversação, à medida que interagia com os humanos, imitando assim os seus padrões de fala. O público-alvo deste *chatbot* eram jovens entre os 18 e 24 anos e foi criada de forma a estar preparada para potenciais ataques e abusos, com a implementação de

filtragem e estudos aprofundados de utilizadores a fim de providenciar uma boa experiência (Lee P. , 2016).



Figura 2. Um dos primeiros *posts* inofensivos de Tay.

Inicialmente os posts de Tay eram normais e inofensivos (Figura 2) mas em pouco tempo ficou exposta e totalmente vulnerável a ataques de conteúdo ofensivo, quando os utilizadores interagiam com esta. Esses ataques, provenientes na sua maioria pela comunidade do 4chan<sup>3</sup>, conduziram entre outros, a posts ofensivos nazis, anti-feministas, racistas (Figura 3). Por conseguinte, os comentários mais ofensivos começaram a ser eliminados, mas não sendo o suficiente, a Microsoft decidiu colocar o *bot* indisponível em menos de 24 horas de existência (Hern, 2016).



Figura 3. Exemplo de um post ofensivo de Tay.

Depois desta tentativa, a Microsoft tem-se deparado com certos desafios de pesquisa no *design* da Inteligência Artificial e tentado representar o melhor da humanidade mesmo não podendo prever todos os ataques abusivos sem deparar-se com os erros (Lee P. , 2016).

<sup>3</sup> Forum baseado na partilha e discussões de imagens mantendo o anonimato.

### 2.1.2.2. Zo

No fim de 2016 a Microsoft lançou o *chatbot* Zo baseado na tecnologia de Xiaoce e Rinna, *chatbots* da Microsoft de sucesso na China e Japão. Zo está disponível em plataformas de *chat* como o Facebook Messenger e o Kik. Assim como Tay, também tem como finalidade simular uma jovem humana. Zo distingue-se de Tay na medida em que é composta por uma filtragem mais eficaz de conteúdos. Essa filtragem é notável, pois se um utilizador abordar por exemplo o tema de política, Zo responde pedindo para alterar o assunto da conversa (Figura 4).

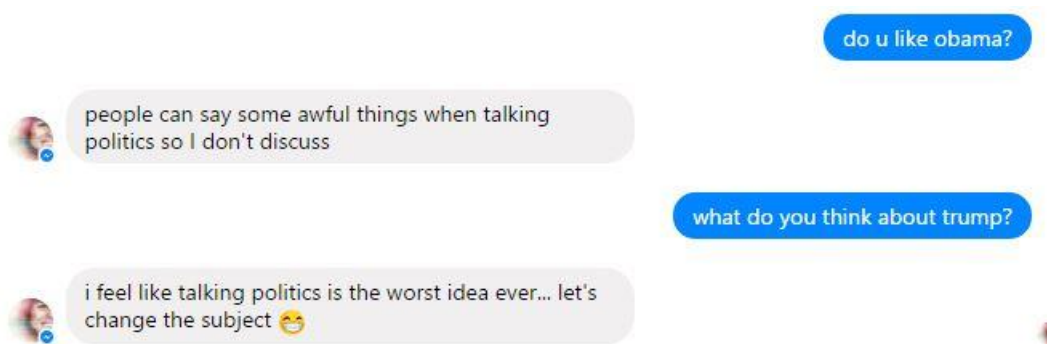


Figura 4. Zo tem capacidade de evitar certos temas como a Política.

Este *chatbot* tem iniciativa própria, no caso de um utilizador ter interagido anteriormente, pois num momento aleatório tenta interagir novamente com o utilizador.

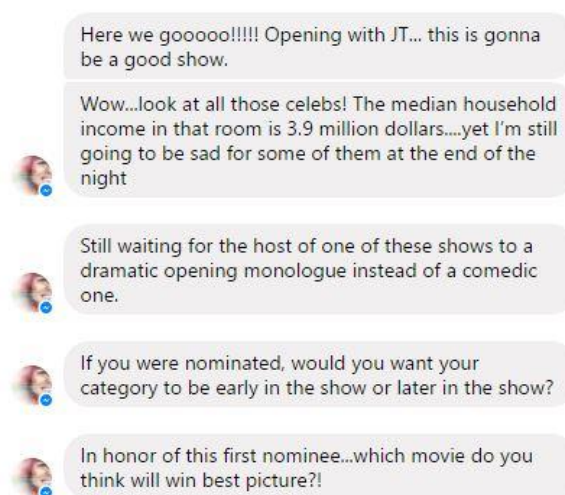


Figura 5. Iniciativa própria de Zo para interagir com o utilizador.



Por exemplo, na noite de Óscares 2017, Zo teve a iniciativa de convidar os utilizadores a assistirem a gala juntamente com ela. Caso o utilizador aceitasse o convite através do comando “#Awards”, Zo comentava os Óscares durante a noite, tentando manter a interação com o utilizador (Figura 5).

### 2.1.3. Assistentes de Voz

Os assistentes de voz são agentes que funcionam como agentes pessoais. Assim como os *chatbots*, estes assistentes interagem através da linguagem natural. No entanto, ao contrário dos *chatbots* que só funcionam com entradas e saídas de texto, os assistentes de voz são capazes de compreender a entrada do utilizador de linguagem natural, maioritariamente através de voz, ou em alguns casos também texto. Um assistente de voz é vantajoso pois permite ao utilizador interagir e obter respostas através da voz, podendo assim realizar outras tarefas ao mesmo tempo, tendo em conta que as suas mãos não ficam ocupadas para interagir com o sistema. No entanto, a interação por voz pode não ser adequada para determinados contextos, como locais públicos, onde pode estar presente um grande ruído.

Grandes empresas de tecnologia recorrem à criação de assistentes de voz, como é o caso da Cortana da Microsoft, a Siri da Apple, o Google Assistant da Google e a Alexa da Amazon.

Assistentes como estes podem ter uma interação personalizada com o utilizador, reproduzir música, adicionar lembretes, enviar e fazer chamadas para alguém, entre muitos outros.

Nos próximos subcapítulos (2.1.3.1, 2.1.3.2 ) serão destacados vários assistentes de voz com a finalidade de analisar as suas funcionalidades e fazer uma comparação entre todos.

#### 2.1.3.1. Cortana

A Cortana, criada pela Microsoft em 2014, funciona como uma assistente pessoal para o utilizador pois é capaz dar alertas, informação sobre pesquisas e até mesmo manter um diálogo. A Cortana pode interagir de forma personalizada pois caso o utilizador tenha sessão iniciada, esta assistente pode ter acesso a dados necessários como é o caso dos contatos e até sugestões relacionadas com o perfil do utilizador. Para os utilizadores que não têm sessão, a assistente continua a ter capacidade para fazer inúmeras coisas como pesquisas, colocar alarmes, responder a questões, entre outros, mas sem personalização.

A Cortana é capaz de realizar várias tarefas com o dispositivo bloqueado e é possível ativa-la com a entrada de voz “Hey Cortana”. Após esta entrada, a Cortana fica ativa, recolhendo a restante entrada do utilizador depois de ouvir as palavras-chave.

Esta assistente também está apta para conectar-se ao telemóvel com a sessão iniciada no computador e no telemóvel com a conta da Microsoft. Assim, a Cortana pode alertar o utilizador de chamadas perdidas e enviar mensagens através do computador (Microsoft).

### 2.1.3.2. Alexa – Amazon Echo

Alexa, surgida em 2014, é a assistente pessoal criada pela Amazon e está armazenada em vários dispositivos como o Amazon Echo, Echo Dot, e Amazon Tap. Esta assistente é composta por mais do que 10.000 *skills*<sup>4</sup> criadas por programadores, sendo capaz de definir alarmes, alterar a temperatura da casa, desligar/ligar luzes, reproduzir música, reservar serviços, comprar produtos, ler notícias, responder a perguntas de diversas áreas, entre outras.

Alexa tem a capacidade de armazenar informação de interações anteriores, recorrendo aos interesses e preferências do utilizador, adaptando-se a este e ao seu vocabulário.

O primeiro dispositivo e o principal, que surgiu em 2015, é o Amazon Echo. A interação baseia-se nas entradas e saídas de voz, tendo apenas três *botões* físicos para ligar e desligar o Echo, desligar o microfone e controlar o volume. Está apto para reconhecer voz a altas distâncias pois tem sete microfones que eliminam o ruído do ambiente, podendo ficar situado no centro da casa. O Echo pode ser ativado com a palavra “Alexa”, não sendo necessário recorrer ao toque para tal.

Apesar das diferentes ofertas de dispositivos para a Alexa assim como a melhoria constante destes e de Alexa, com o surgimento do Google Home, a Amazon tem o objetivo de melhorar a linguagem de Alexa, assim como dar-lhe a capacidade de detetar as emoções do utilizador através do seu tom de voz. Existem *softwares* para detetar as emoções do utilizador através da sua voz, como o Vokaturi e Good Vibrations, mas nos últimos anos têm surgido novas formas, ainda com problemas, para as máquinas aprenderem a reconhecer as emoções com maior eficácia (Knight, 2016).

### 2.1.3.3. Google Assistant – Google Home

A Google Assistant, criada em 2016, é a assistente da Google que pode ser encontrada no dispositivo doméstico Google Home. Esta é uma assistente pessoal que tem a capacidade de realizar as mais diversas tarefas. Tal como Alexa, também pode responder a questões de diferentes áreas, fornecer informação sobre o trânsito, controlar dispositivos domésticos e muitas outras.

Esta assistente consegue superar a Alexa em certos pontos. Um deles está relacionado com a sua inteligência pois a Google Assistant tem a capacidade de compreender quando o utilizador cometeu um erro e quer retifica-lo, enquanto Alexa não tem essa capacidade. Esta, ao contrário de Alexa, tem iniciativa própria pois consegue iniciar um diálogo com o utilizador e é capaz de gerar

---

<sup>4</sup> Grupo de habilidades disponíveis na Alexa Skills para o utilizador poder melhorar a sua interação.

respostas contextualizadas. Posto isto, caso o utilizador pergunte “Quantos anos tem Hemma Watson?” e depois perguntar “Em que filmes participou?” sem mencionar o nome da atriz, a Google Assistant consegue identificar que o contexto é o mesmo, enquanto que Alexa apenas diz que não consegue responder a esse pedido.

O dispositivo Google Home, tal como o Amazon Echo baseia-se na interação através de entradas e saídas de voz. Apenas é composto por um painel no qual é possível reproduzir e parar músicas e alterar o volume e um *botão* para desligar o microfone. O Google Home consegue interagir a partir das palavras “OK Google” ou “Hey Google” não sendo necessário ter contacto físico com o dispositivo para iniciar um diálogo.

Este pode ficar situado em qualquer parte da casa conectado a outros dispositivos iguais situados noutras divisões. Também é constituído por microfones capazes de eliminarem ruído ambiente, estando apto para reconhecer a voz do utilizador a grande distância.

#### **2.1.4. Interfaces Human-powered**

Além dos *chatbots* e assistentes de voz, existem outro tipo de interfaces: as *human-powered* em que o agente de conversação não é um programa, mas sim um humano. Certas interfaces deste tipo estão a evoluir em relação à tecnologia utilizada pois estão a recorrer ao uso de inteligência artificial. A inteligência artificial tem o intuito de memorizar e armazenar os dados de cada utilizador para assim a sua interação ser personalizada, apoiando simultaneamente o trabalho do agente humano. Posto isto, o agente humano é capaz de responder mais rápido aos pedidos do utilizador e de acordo com o seu perfil.

De seguida (subcapítulos 2.1.4.1 e 2.1.4.2) serão abordadas duas interfaces que revolucionaram a forma de *e-commerce*, uma vez que permitem efetuar compras através de mensagens com a ajuda de agentes humanos.

##### **2.1.4.1. Magic**

A Magic é uma plataforma surgida em 2015 que integra agentes humanos, inteligência artificial e um *software* para responder e realizar rapidamente as necessidades dos utilizadores. Magic é um serviço composto por um número para qual o utilizador envia mensagens, não necessitando assim de nenhuma aplicação para o seu uso (Figura 6). No entanto, o utilizador tem a alternativa de usufruir dos serviços de Magic com na aplicação Telegram, através de um *chatbot*. Assim, este pode utilizar os serviços através do computador para além do *smartphone*, caso seja mais conveniente.



Figura 6. Processo de compra na plataforma Magic.

Magic satisfaz quase todas as necessidades do utilizador. Pode efetuar compras e reservas, planear viagens, encomendar produtos, agendar eventos no calendário do Google e muitas outras. Quanto mais o utilizador recorre a Magic, mais informação sobre este é guardada recorrendo a inteligência artificial. São guardadas as suas preferências, podendo assim o agente humano ser mais rápido e eficaz nas respostas. Por exemplo, se um utilizador não gostar de uma determinada cor para roupa, a inteligência artificial de Magic automaticamente removerá essa cor dos resultados de roupa pedidos (Magic, 2015).

#### 2.1.4.2. Operator

A Operator, é uma aplicação surgida em 2014 que da mesma forma que Magic, é composta por agentes humanos e inteligência artificial. No entanto, é simulada uma relação mais próxima entre os agentes e os utilizadores, pois é possível ver uma fotografia e o nome dos agentes através das mensagens.

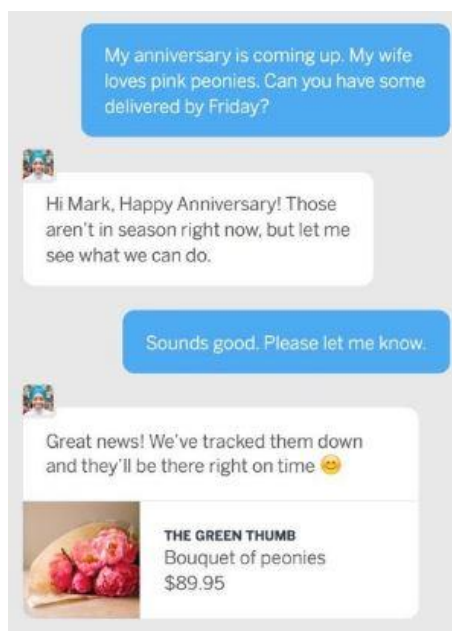


Figura 7. Simulação de proximidade entre o agente e o utilizador.

A Operator tem como finalidade dar a oportunidade ao utilizador de comprar qualquer produto que necessita, através de um serviço personalizado e eficaz, como acontece nas lojas físicas. Os agentes humanos ajudam o utilizador a chegar ao melhor produto de acordo com as suas preferências. Isto é feito através da apresentação de uma lista de produtos com base no que o utilizador solicita. Esta lista de produtos também é criada com a ajuda da inteligência artificial, que tal como na plataforma Magic, também armazena as preferências do utilizador a fim de adquirir respostas rápidas e personalizadas (Operator).

### 2.1.5. Comércio Conversacional

A procura de produtos em plataformas de venda *online* tem vindo a aumentar, no entanto, tem-se notado um abandono do carrinho de compras, onde cerca de 68% das compras *online* não são concretizadas. Alguns dos motivos para esta perda incluem o longo processo que os clientes enfrentam para realizar compras *online*, dúvidas e problemas de usabilidade nos diversos dispositivos (Baymard Institute).

As compras *online* podem ser as preferidas para o consumidor pois são rápidas, encontra-se uma maior variedade de produtos e há uma maior facilidade em comparar preços. Todavia, como já referido anteriormente, este método de realizar compras pode-se tornar mais longo.

De acordo com os dados do BI Intelligence (2016), o número de utilizadores de aplicações de mensagens está a aumentar, comparativamente às redes sociais (Figura 8).

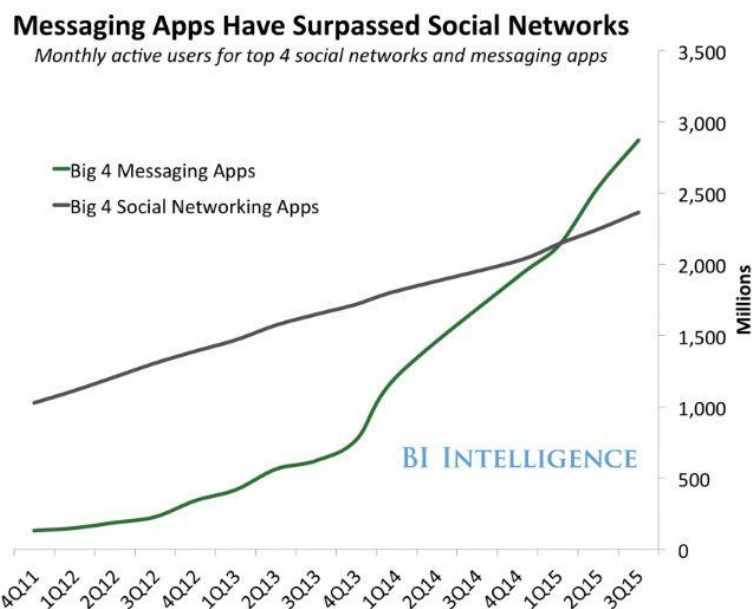


Figura 8. Utilização crescente das aplicações de mensagens (BI Intelligence , 2016).

Essas aplicações de mensagens – *chats* - têm apresentado uma constante evolução e uma maior utilização para os utilizadores comunicarem com marcas e procura de outros conteúdos, para além de continuarem a comunicar com amigos. Devido a várias características dos utilizadores destas plataformas, particularmente o facto de serem jovens e as usarem com muita frequência, tornaram-se grandes focos de interesse para as marcas introduzirem um ponto de negócio. Além de existirem marcas que investem na criação de plataformas como estas para aumentar o seu negócio, outras resumem-se unicamente à criação de uma plataforma para si mesmas, como é o caso da Magic.

As plataformas de comércio conversacional permitem ao utilizador comunicar com *bots*, agentes humanos, serviços ou comunicar de forma híbrida - com agentes humanos e inteligência artificial em simultâneo. O pedido de serviços ou compras podem ser realizados através de plataformas de *chat*, como o Facebook Messenger, através de mensagens, como o Operator e Magic, ou através de dispositivos principalmente direccionados para reconhecimento de voz, como é o caso do Amazon Echo.

Messina (2016), criador do termo “comércio conversacional<sup>5</sup>”, afirmou que 2016 seria o ano do deste tipo de comércio. Esta forma de realizar compras tornou-se popular em 2015 quando a Uber se conectou ao Facebook Messenger, permitindo aos clientes requisitar os seus serviços. Desde então, muitas empresas estão a investir na criação de *chatbots* para alargarem a sua marca, permitindo assim que os clientes tenham oportunidade de realizar compras com maior facilidade (Messina, 2016).

<sup>5</sup> Tradução de *Conversational Commerce* (Messina, 2016)

Com estas plataformas de *chat*, as empresas podem comunicar facilmente com os utilizadores e otimizar a experiência de compra, não sendo necessário para o utilizador recorrer a outra aplicação para realizar compras.

## 2.2. Design de Usabilidade

Nos seguintes subcapítulos (2.2.1, 2.2.2 e 2.2.3) são abordados vários pontos a ter em conta para a realização do *design* de uma interface ser capaz de otimizar a experiência do utilizador.

### 2.2.1. Interação Humano-Computador

A IHC foca-se na forma como ocorre a interação entre o humano e a máquina assim como os aspetos tecnológicos que influenciam essa interação. Estando baseada na Ergonomia e nos Fatores Humanos, a IHC tem então como finalidade providenciar uma boa experiência ao utilizador independentemente do sistema em causa, tanto em relação a questões físicas como questões cognitivas e psicológicas.

O termo IHC não se limita a representar um utilizador e um computador. Neste aspeto o humano simboliza qualquer utilizador ou grupo de utilizadores que realizam determinada tarefa, recorrendo a alguma tecnologia e por computador entende-se qualquer tipo de tecnologia. Por fim, interação refere-se à comunicação que ocorre entre o humano e o computador.

A IHC envolve diferentes áreas como a Ergonomia que se relaciona com as capacidades físicas, *Design* Gráfico para a interface ter um visual eficaz, Sociologia para compreender o contexto da interação, Ciência da Computação e Engenharia para construir a tecnologia necessária, entre outras. No entanto, a Ciência da Computação e o *Design* do sistema são os principais focos, originado que a IHC seja composta por três fatores: conceção, implementação e avaliação dos sistemas com base nas tarefas do utilizador.

Apesar de a IHC estar focada essencialmente em compreender a forma como os humanos interagem com a tecnologia, o *Design* também apresenta um papel fundamental na IHC. Foi definida uma regra muito importante de *design* que passa por compreender os componentes em causa.

Para realizar o *design* de uma interface é necessário compreender os componentes essenciais na IHC: o humano e o computador. Para o humano é necessário compreender os erros que pode cometer durante a interação que potencialmente conduzem a erros do sistema, a psicologia humana e os aspetos sociais. Em relação ao computador, é necessário compreender as suas capacidades e limites de acordo com a tecnologia, conhecer as plataformas que existem de forma a analisa-las e as suas ferramentas (Dix, Finlay, Abowd, & Beale, 2004).

### 2.2.2. Usabilidade

De acordo com Nielsen (2012), a usabilidade é responsável por determinar a facilidade com que o utilizador se depara ao interagir com uma interface. Esta, juntamente com a utilidade apresentam o mesmo grau de importância, pois determinam o que pode ser útil para o utilizador. Desta forma, a facilidade de interação perde importância se a interface não corresponder às necessidades do utilizador.

Unger e Chandler (2012) defendem que para conceber experiências memoráveis, o *UX Designer* necessita de compreender como pode criar uma harmonia e estrutura viável, assim como quais os elementos importantes para existir uma ligação emocional com os utilizadores.

Para corrigir uma aplicação é fundamental que sejam realizados testes de usabilidade e estes devem ser produzidos nas fases iniciais do projeto (Cuello & Vittone, 2013). Estes testes ajudam na correção e melhoria de uma aplicação, com base em observações dos utilizadores. Nielsen (2012) considera que para medir a usabilidade de uma interface é necessário ter em consideração cinco atributos:

1. Fácil aprendizagem: se a interface for de fácil aprendizagem, o utilizador realiza as tarefas pretendidas mesmo na sua primeira utilização, aprendendo rapidamente como interagir com a mesma;
2. Eficácia: a eficiência da interface depende do nível de eficácia com que o utilizador interage com esta, após ter aprendido como a interface funciona.
3. Fácil memorização: se o sistema for simples de memorizar, o utilizador facilmente se lembra como interagir com a interface, mesmo após algum tempo desde a sua última interação.
4. Erros: o sistema deve ter uma baixa probabilidade de erros e devem ser de fácil recuperação.
5. Satisfação: um dos objetivos principais em apostar na usabilidade do sistema, é ter como resultado a satisfação e prazer por parte do utilizador ao ter contacto com o *design* da interface.

Existem vários métodos para estudar a usabilidade, no entanto Nielsen (2012) assume que o método mais útil é o teste do utilizador, composto por três componentes: conhecer os utilizadores que representam o público-alvo, pedir que os utilizadores realizem as tarefas representativas e por último, observar as ações dos utilizadores, tendo em conta todos os passos e resultados das tarefas realizadas.

Apesar da presença de vários processos para atingir uma interface de utilizador de alta qualidade, Nielsen (2012) considera que o processo mais eficaz e económico é o *Design Iterativo* que consiste num ciclo de repetição das diferentes fases: desenhar, testar e analisar.

Para uma potencial melhoria da usabilidade e utilidade do sistema, assim como o aumento de vontade em utilizá-lo, pode-se recorrer às Dez Heurísticas de Usabilidade de Nielsen e Molich (Nielsen J. , 1995):



1. Visibilidade do estado do sistema: o sistema deve transparecer o seu estado ao utilizador, isto é, deve informar o utilizador das operações que estão a decorrer durante um determinado tempo;
2. Correspondência entre o sistema e o mundo real: o *designer* deve facilitar a experiência do utilizador, tendo consideração por aspetos cognitivos, imitando conceitos utilizados numa experiência real;
3. Controlo e liberdade do utilizador: providenciar aos utilizadores a oportunidade de retroceder e refazer ações;
4. Consistência e padrões: é necessário manter a interface coesa em relação aos elementos e as respetivas funcionalidades;
5. Prevenção de erros: é necessário reduzir ou prevenir o número de erros, podendo eliminá-los ou prevenindo com um aviso as ações que podem levar a erros;
6. Reconhecer em vez de recordar: devido à memória de curto prazo dos humanos, é necessário ter em consideração que o humano só consegue memorizar cerca de cinco itens ao mesmo tempo. Em consequência, há que optar pelo reconhecimento que pode ser considerado mais fácil do que relembrar algo;
7. Flexibilidade e eficiência de uso: o utilizador deve ter a oportunidade de alterar e adaptar a interface de acordo com as suas necessidades, conduzindo a uma maior facilidade de utilização para ele;
8. Design estético e minimalista: apresentar só o conteúdo útil é vantajoso para que o utilizador não tenha distrações com conteúdos de pouca importância, resultando num melhor controlo da capacidade de memória de curto prazo dos humanos;
9. Ajudar o utilizador a reconhecer, diagnosticar e recuperar dos erros: o sistema deve transmitir de forma clara as mensagens de erros para o utilizador ter a certeza do que está a ocorrer;
10. Ajuda e documentação: o utilizador pode necessitar de recorrer a documentação para tirar certas dúvidas, sendo importante chegar à informação facilmente e apresentá-la de forma clara.

Ben Shneiderman (1997) apresentou princípios heurísticos dominados por oito Regras de Ouro. Essas regras são semelhantes às Dez Heurísticas de Usabilidade de Nielsen e Molich e também são para ter em consideração na conceção do *design* de interfaces, com o objetivo de melhorar a usabilidade e *design* de interação:

1. Consistência: a coesão é importante para os utilizadores ficarem familiarizados com a interface e como a informação é transmitida;
2. Fornecer atalhos para os utilizadores frequentes: para casos de uso frequentes, o ideal é existirem atalhos ou outras formas de facilitar a interação para que o utilizador possa interagir com mais rapidez;

3. Feedback informativo: o utilizador deve ser informado do que está a suceder na interface de forma apropriada e simples durante um determinado período de tempo;
4. Design de diálogo para fechar: o utilizador deve ser notificado quando termina uma ação importante com o intuito de tomar conhecimento e preparar-se para as próximas ações;
5. Oportunidade de recuperação de erros simples: é necessário o utilizador receber instruções de como recuperar dos erros que ocorrem de forma clara e rápida;
6. Retroceder facilmente de ações: o utilizador deve ter a opção de retroceder as suas ações. Segundo Shneiderman essa oportunidade encoraja os utilizadores a explorarem o sistema sem receio;
7. Suporte local interno de controlo: dar a oportunidade ao utilizador de iniciar a ação, parecendo que estão a controlar a interface;
8. Reduzir carga de memória de curto prazo: a capacidade humana de processar informação de curto prazo cria a necessidade de apresentar a informação de forma sucinta e simples.

Existe a probabilidade de se pensar que o principal objetivo e dificuldade, essencialmente nas fases iterativas, é encontrar os problemas de usabilidade para assim resolvê-los. No entanto, a maior dificuldade no *design* da interface não passa por encontrar os problemas em relação à usabilidade e corrigi-los, mas sim perceber quais os problemas que valem a pena serem corrigidos, a fim de não gastar tempo desnecessário (Dix, Finlay, Abowd, & Beale, 2004)

### 2.2.3. Protótipos

O protótipo é o processo onde são realizados *mockups* da interface, de baixa e alta qualidade, a fim de poder testar esta mesma interface com os utilizadores. Estes são uma boa forma de visualizar e testar o *design* de uma forma concreta, sem recorrer à programação e implementação, permitindo ao *designer* centrar-se unicamente na interação do utilizador com o sistema. De acordo com Billy Hollis, nos testes dos protótipos apenas devem ser avaliados critérios qualitativos e não quantitativos, para que o objetivo não se afaste do mais importante – a interação do utilizador com o sistema (Krug, 2000).

Recorrer a testes de protótipos é uma boa forma de auxiliar o *designer* para aprender sobre as interações dos utilizadores e assim melhorar vários pontos do *design*. Existem diversos tipos de protótipos para as mais variadas necessidades: página única, ao invés de várias páginas com as funcionalidades necessárias para realizar as tarefas, protótipos realistas de alta fidelidade em oposição aos esboços de baixa fidelidade e protótipos interativos (em que utilizador pode interagir sozinho com o sistema), em vez de estáticos (onde uma pessoa é que realiza as ações pelo computador (Pernice, 2016).

O tipo de protótipo a escolher está dependente de vários componentes, como os objetivos do próprio teste, as ferramentas utilizadas para a conceção do protótipo assim como os meios que estão a cargo de poder ajudar antes e durante os testes de usabilidade (Pernice, 2016).

Num protótipo interativo, o objetivo passa por o *designer* definir, previamente ao teste, como o sistema vai reagir consoante a interação do utilizador.

No protótipo estático, não é definida nem implementada nenhuma resposta do sistema porque estas vão ser em tempo real, ou seja, durante o teste, o *designer* responderá no lugar do sistema. Para este fim, existem diferentes métodos que podem ser utilizados como o WoZ (*Wizard of Oz*) e o protótipo de papel.

No método WoZ, o *designer* – *Wizard* – enquanto analisa as ações do utilizador a partir de outra sala, tem controlo sob a interface, gerando as respostas do sistema de acordo com as interações do utilizador. Visto que o sistema é estático, nenhuma interação do utilizador tem qualquer influência no protótipo, por isso o *designer* responde para o utilizador não notar que o sistema é estático.

Este método inicialmente era utilizado para testar interfaces de linguagem natural, mas ao longo dos anos tem sido utilizado noutro tipo de interfaces (Beaudouin-Lafon & Mackay, 2003). Ainda assim, este género de teste é considerado útil para testar sistemas, antes de serem implementados, que funcionam com inteligência artificial (Pernice, 2016).

O protótipo realizado em papel, à semelhança ao WoZ, também apresenta as respostas criadas pelo *designer*. Neste teste, o *designer* tem posse dos papéis, onde está presente o *design* da interface, e de acordo com as ações e gestos do utilizador, o *designer* coloca a folha com a determinada resposta do sistema.

## 2.3. Design em Interfaces Conversacionais

Ao longo dos anos tem sido evidente um maior destaque na usabilidade das interfaces conversacionais. Porém, não têm sido utilizados tantos recursos na usabilidade, quando comparada com a tecnologia. Mesmo assim, a usabilidade agora é vista como um elemento competitivo de alta importância que pode ser a chave para a eficiência e qualidade da interface conversacional, assim como para uma maior satisfação por parte do utilizador (Dybkjaer & Bernsen, 2000).

Ao contrário das interfaces gráficas que podem mostrar visualmente funcionalidades e informações ocultas, as interfaces conversacionais são mais complexas de se realizar, uma vez que apresentam pouco ou mesmo nenhum conteúdo visual. Como não possuem uma forma de apresentar visualmente funcionalidades como menus e opções, é necessário recorrer a outras técnicas para uma melhor interação (Yankelovich, 1996).

Um problema comum para a usabilidade neste tipo de interfaces ocorre quando os utilizadores não têm conhecimento dos limites de reconhecimento da interface, podendo dizer algo que a

interface não está apta para reconhecer. Uma solução para o problema em causa passa pela criação de uma *prompt* que apresente as opções disponíveis na interface. No entanto, devido às capacidades de memória de curto prazo dos humanos, se a *prompt* for extensa, os utilizadores provavelmente só se vão recordar da informação apresentada no fim desta (Yankelovich, 1996). Outra solução, tendo em conta a memória auditiva e os seus limites, passa pelo sistema apresentar um sumário quando necessário (Möller, 2005).

Deste modo, a interação do utilizador com a interface fica dependente de fatores humanos como a memória de curto prazo, atenção, clareza, dicção, audição e até o ruído do ambiente (Schnelle-Walka & Lyardet, 2006).

É importante uma interface conversacional ser composta por determinadas propriedades como a independência do utilizador, a fala contínua e os vocabulários grandes, que podem influenciar positivamente a experiência do utilizador. A independência do utilizador permite que o sistema admita e reconheça com grande precisão a entrada de vários utilizadores. Este aspeto é importante pois o sistema passa a ser acessível aos utilizadores ocasionais. Quando o utilizador necessita de dar entradas com as palavras separadas, a sua interação torna-se pouco natural e afeta o desempenho das tarefas. Em consequência, a interface deve ser composta pelo reconhecimento contínuo de fala permitindo que o sistema reconheça palavras ligadas, levando a que o utilizador comunique de forma natural. Em adição, se a interface for composta por um vocabulário grande, também melhora a interação, uma vez que o sistema é capaz de reconhecer com alta precisão as entradas do utilizador.

Por norma, o utilizador concentra-se na tarefa do momento numa situação natural. No entanto, o utilizador pode ter um discurso espontâneo devido a distrações, hesitações e pausas. Por conseguinte, a interface não deve restringir palavras e é necessário estar preparada para termos fora do contexto (Cohen, Giangola, & Balogh, 2004).

As *prompts* podem ter um grande impacto na opinião do utilizador no que diz respeito à interface. Assim foram criados certos pontos a ter em conta para o *design* das mesmas. Estas devem ser breves, para não massacrar os utilizadores com demasiada informação, apresentar linguagem coesa e uma voz agradável face aos humanos. Em caso de necessidade de repetir a *prompts*, se o utilizador não se recordar de certa informação, é fundamental fornecer mais informações na repetição, de forma a ajudar o utilizador a compreender de forma eficaz (Fraser, 1997). Contudo, mesmo seguindo estes pontos para a criação de *prompts*, podem não ser o suficiente para agradar os utilizadores. Existem certos momentos em que não são necessárias as *prompts* para apresentar informação, podendo ser substituídas, por exemplo por *earcons*<sup>6</sup> (Möller, 2005). Consequentemente, a interação poderá ser mais dinâmica e agradável para o utilizador, pois a interface não é monótona.

---

<sup>6</sup> Som que simboliza uma determinada ação no sistema

Outro ponto que influencia a usabilidade e interação de um sistema é o tipo de iniciativa escolhida para a interface. Enquanto que Nielson (1999) defende que o humano é que deve ter o controle sob o sistema, ou seja, iniciativa do utilizador, Tannenbaum (1998) defende que a interatividade apenas acontece quando o humano e a máquina têm a mesma quantidade de controle, ou seja, quando a iniciativa é mista. Em adição, as interfaces que restringem as entradas do utilizador – interfaces com iniciativa do sistema - podem induzir à frustração do mesmo devido à sua limitação de vocabulário e tarefas. Por conseguinte, um sistema com iniciativa mista é o que proporciona uma interação eficiente, pois o utilizador tem controle do diálogo e o sistema tem a capacidade de reconhecer as necessidades do utilizador, dando respostas a estas (Allen, D. Byron, Ferguson, Galescu, & Stent, 2000).

É importante ter em conta os contextos de utilização e ambientes em que a interface será utilizada para tomar decisões mais indicadas quanto às modalidades de interação. Se o utilizador recorrer à interface num sítio público e com grande ruído, não é apropriado fazer a sua entrada no sistema através de voz, mas sim manualmente (Möller, 2005). Deste modo, a melhor solução passa por se optar por uma interface híbrida. Rudnicky & Hauptmann (1989) também apoiam as interfaces híbridas afirmando que a criação destas é um grande objetivo pois é uma forma de tornar a comunicação com o sistema completa.

Recorrendo a uma interface híbrida, os melhores pontos de uma modalidade podem recompensar os pontos mais fracos da outra (Cohen, 1992).

Uma interface híbrida pode ser vantajosa no caso da recuperação de erros. Rudnicky e Hauptmann (1989) realizaram testes com a finalidade de analisar qual seria a melhor forma de correção dependendo do tipo de entrada. Com base nos resultados de três métodos testados, concluiu-se que o tempo de entrada através da fala é uma melhor opção face à escrita. No entanto, a fala tem diversas desvantagens nos textos mais longos, dado que precisam de mais correções. Relativamente ao tempo de correção, esta não é a melhor opção, pois a entrada manual quase não necessita de correções. No momento de correção e confirmação, a escrita é a mais rápida e a fala é mais lenta. Contudo, o tempo de confirmação aumenta dependendo do tamanho da entrada (Rudnicky & Hauptmann, 1989).

Para interfaces interativas eficazes é necessário ter em conta a IHC e a sua natureza multidisciplinar (Dix, Finlay, Abowd, & Beale, 2004). Para proporcionar uma boa experiência ao utilizador, é necessário em primeiro lugar satisfazer as suas necessidades, mas utilizando uma multidisciplinaridade, não dando apenas o que o utilizador quer. Com esta fusão torna-se possível alcançar uma experiência de grande qualidade (Nielsen & Norman).

Posto isto, várias áreas de conhecimento podem ser úteis para apoiar o processo de *design* de uma interface conversacional. Entender as capacidades cognitivas humanas, o que os utilizadores esperam da interação com a interface, compreender os princípios do *design* centrado no utilizador a fim de satisfazer as necessidades do negócio e do utilizador e por fim ter uma base de

conhecimento relativamente à tecnologia da fala, para assim tomar decisões mais acertadas em relação ao *design* (Cohen, Giangola, & Balogh, 2004).

Compreender a tecnologia de fala é importante para tomar decisões de *design* nas interfaces, porque se o *designer* compreender o que a tecnologia afeta em relação ao desempenho, pode utilizar a tecnologia a favor dos pontos fracos de *design*. Durante a interação, o utilizador pode deparar-se com diferentes erros. Assim se existir previsão de potenciais lapsos, é possível criar estratégias harmoniosas para recuperar destes. O *designer* necessita de compreender e ter conhecimento dos impactos no reconhecimento e no processo para assim criar e modificar os ficheiros da gramática e dicionário do sistema consoante o necessário.

O *designer* será capaz de tomar facilmente decisões se estiver envolvido em todas as fases: definição de requisitos, *design* de alto nível, *design* detalhado, desenvolvimento, testes e sintonização. Estas mesmas fases podem ser divididas em três objetivos. As fases de definição de requisitos e *design* de alto nível têm como finalidade a definição, isto é, compreender as necessidades do que envolve aplicação (o utilizador e o negócio) e assim criar um *design* que possa satisfazer essas necessidades. A fase de *design* de alto nível tem como objetivo o *design*, ou seja, aprofundar ao máximo o *design* já realizado na fase anterior. Por fim, a fase de desenvolvimento, testes e sintonização têm a finalidade de realização, significando que é nesta fase que o *design* é implementado no sistema (Cohen, Giangola, & Balogh, 2004).

A avaliação é fundamental nos sistemas conversacionais para os utilizadores assim como para quem desenvolve. Hirschman e Thompson (1997) definem três tipos de avaliações com objetivos diferentes:

1. Avaliação diagnóstica: o seu objetivo é avaliar e encontrar erros de implementação e de *design*, podendo ser utilizada por quem desenvolve o sistema, assim como pelos futuros utilizadores finais. São realizados testes com diferentes contextos e com a hipótese de estarem incluídas entradas marcadas como válidas ou inválidas;

2. Avaliação de desempenho: tem a finalidade de medir o desempenho do sistema em várias áreas específicas recorrendo a parâmetros qualitativos e quantitativos. Tendo em conta a oportunidade de avaliação de desempenho, é necessária uma distinção entre níveis: o critério de interesse a ser avaliado como a precisão, taxa de erro e velocidade; a medida, isto é, o atributo a ter em conta para o critério escolhido, como o tempo de processo, rácio de acessos e erros; e o método que define a forma como é determinado o valor para uma certa medida e sistema.

3. Avaliação de adequação: tem o intuito de avaliar se o sistema é capaz de corresponder sem problemas à finalidade para que foi criado, satisfazer as expectativas e necessidades do utilizador e se é o mais adequado comparando com os outros existentes. Todavia, podem ser necessárias avaliações de desempenho e diagnóstico para fornecer ao utilizador a comparação entre os sistemas semelhantes a fim de o utilizador fazer uma escolha com bases (Hirschman & Thompson, 1997).

Os aspetos de *design* de uma interface conversacional afastam-se da tecnologia do reconhecimento de fala e focam-se no modo de execução de uma tarefa. Rudnicky & Hauptmann (1989) definiram seis princípios básicos de *design* de interação para estas interfaces, mas que também podem ser utilizados noutro tipo de interfaces, que resultaram da experiência com sistemas de fala desenvolvidos e da investigação empírica na interação de fala:

1. Flexibilidade do utilizador: os humanos são flexíveis na interação pois têm a capacidade de adaptarem a sua fala ao sistema com que interagem, sendo já comum os humanos alterarem o seu tipo de fala de acordo com o contexto. Enquanto que num diálogo entre pessoas é notável um desleixo no discurso, num diálogo com uma máquina o humano naturalmente tenta realizar um discurso direto ao assunto e de forma ordenada

2. Estilos de protocolo de interação: as interfaces de voz nem sempre detetam de forma correta as entradas do utilizador, no entanto, com protocolos de interação é possível controlar a interação dos participantes fazendo correções e confirmações.

O primeiro protocolo (Figura 9 – adaptada de Figura 9) apresenta pouca usabilidade para o utilizador pois se o sistema não reconhecer a entrada ou se produzir uma mensagem de erro, o utilizador necessita de desfazer ações ou repetir a entrada.

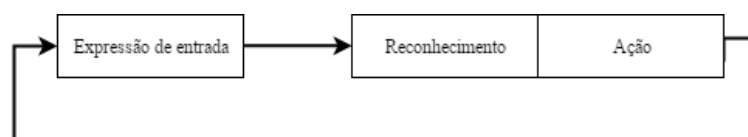


Figura 9. Primeiro protocolo de interação

Ao primeiro protocolo foi adicionado um novo passo – confirmação de entradas – tendo assim um segundo protocolo (Figura 10 – adaptada de (Rudnicky & Hauptmann, 1989)). O utilizador necessita de confirmar o reconhecimento da sua entrada e só assim o sistema pode continuar a acção. A maior desvantagem é a repetição do passo pois é necessário confirmar o reconhecimento em cada entrada.

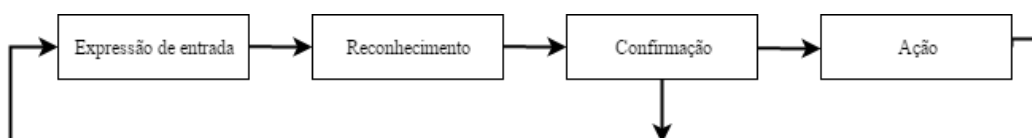


Figura 10. Segundo protocolo de interação

No terceiro protocolo (Figura 11 – adaptada de (Rudnicky & Hauptmann, 1989)) ainda é acrescentado um passo em relação ao segundo. Para além de poder aceitar ou rejeitar o reconhecimento, também está apto de corrigir a entrada

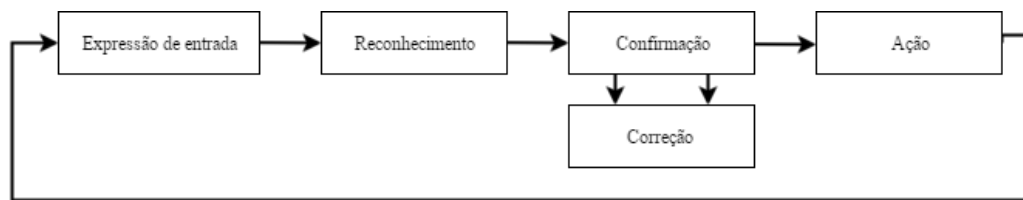


Figura 11. Terceiro protocolo de interação

Segundo o estudo de Rudnicky & Hauptmann (1989) a correção de erros de reconhecimento através do teclado reduz 21% do tempo da tarefa. É possível corrigir através da fala mas se existir uma oportunidade híbrida recorrendo ao teclado e à fala, estes sistemas tornam-se mais eficazes.

O protocolo a ser utilizado depende dos requisitos da tarefa que será executada assim como as características de resposta do reconhecimento. O primeiro protocolo é indicado para sistemas com alta precisão de reconhecimento, enquanto os outros dois são mais adequados para sistemas com baixa precisão de reconhecimento e que necessitam de entradas corretas (Rudnicky & Hauptmann, 1989).

3. Facilidades de correção: a usabilidade de um sistema está dependente da simplicidade e eficácia para confirmar as entradas, mas também da rapidez e eficiência de correção dos erros de reconhecimento de fala.

Existem várias estratégias para o utilizador reparar erros de reconhecimento. Este pode repetir para o sistema, o que não foi compreendido, tal como acontece num diálogo entre humanos. Essa repetição pode ser dada com ênfase, esperando que o sistema reconheça as palavras que deram erro. O utilizador também tem a opção de corrigir conforme o contexto, através da repetição do conjunto de palavras onde se situa o erro. O sistema ao apresentar uma lista de palavras semelhantes às que identificou na entrada, permite ao utilizador escolher a opção correta. Existem outras estratégias através da fala, mas corrigir através da escrita é melhor do que repetir a entrada, uma vez que o mais provável é que o sistema não tenha reconhecido certas partes e só essas precisam de ser editadas.

4. Tempo de resposta: os atrasos no reconhecimento de fala podem afetar a usabilidade. O tempo de resposta de um sistema é essencial para uma boa interação e pode ser decisivo na continuação da tarefa se for muito longo, afectando a usabilidade do sistema.

5. Estrutura de diálogo específica de tarefas: investigar a fundo a estrutura de interação que permite ao sistema prever o que utilizador pode dizer, reduzindo a complexidade dos problemas de reconhecimento de fala.

Normalmente a utilização de um sistema de fala realiza-se no contexto de uma tarefa. Ao ter conhecimento dessa e como as pessoas a realizam, pode ajudar a criar restrições para o reconhecimento de fala. Ao reduzir o número de palavras, a precisão de reconhecimento melhora. Por conseguinte, o desempenho do sistema também melhora e tarefas mais complexas podem ser implementadas.



6. Interação multimodal: é importante avaliar as oportunidades que o utilizador pode ter para interagir de formas diferentes (voz, toque, gestos...) com a interface, consoante o mais adequado no momento.

Gould e Lewis (1985) também recomendam três princípios de *design* desde a década de 1970. Estes podem ser óbvios mas nem sempre são utilizados nos projetos:

1. Foco inicial nos utilizadores e tarefas: os *designers* devem estudar o comportamento e capacidades cognitivas do utilizador para assim o compreender melhor.
2. Medição com base na experiência: os utilizadores são submetidos a testes com protótipos para serem analisadas as suas reacções e como interagem com a interface, de forma detetar eventuais deficiências de usabilidade, assim como se terminam as tarefas eficazmente.
3. Design iterativo/repetitivo: o projeto deve conter um ciclo repetitivo, onde existe a fase de testes, a sua análise e a correcção de erros. Assim, após os testes realizados na fase anterior, os erros detetados necessitam de ser corrigidos. Devido à opinião de que estes princípios são óbvios, Gould e Lewis (1985) fizeram um inquérito a fim de provar que os seus princípios nem sempre são utilizados. Os resultados do inquérito mostram que os seus princípios apenas são óbvios quando são apresentados, pois os *designers* não os compreendem totalmente e são raramente postos em prática.

## 2.4. Análise de Aplicações Relacionadas

Após uma análise ao estado de arte de interfaces conversacionais no contexto geral, tanto como em interfaces conversacionais relacionadas com o comércio, foi necessária uma análise particularmente na área de *design* de usabilidade para este tipo de interfaces. Com esta última análise, surgiu a necessidade de analisar aplicações *e-commerce*, mesmo não sendo conversacionais, para compreender como as interfaces de comércio funcionam a respeito de usabilidade assim como funcionalidades.

As aplicações de *e-commerce* escolhidas para serem analisadas, foram escolhidas com base na sua popularidade internacional para a realização de compras de retalho e são elas: Google Express, Instacart, FreshDirect e Shipt.

Esta análise tem como objetivo compará-las, tendo como base características tais como a página inicial, organização desde as categorias até ao produto, rapidez e facilidade em adicionar e eliminar produtos no carrinho de compra. Ainda existem outras características a analisar que estão diretamente relacionadas com os objetivos da plataforma Shelf AI. Uma das finalidades da Shelf AI passa por fornecer ao utilizador uma experiência de compra fácil, rápida e natural. Por conseguinte, serão avaliados outros aspetos nestas aplicações para analisar que métodos de

pesquisa estão disponíveis, a possibilidade de procurar vários produtos em simultâneo, a compatibilidade dos resultados com a semântica, assim como a oportunidade de filtrar a pesquisa.

### 2.4.1. Google Express

A Google Express, nasceu em 2013 e consiste numa aplicação que permite a realização de compras em diversas áreas. Ela fornece a opção de criar uma lista de compras e a partir desta passar diretamente para os resultados dos produtos da lista para poder comprá-los. Esses produtos na lista de compras também podem ser adicionados diretamente através da página do produto. Também tem a uma secção onde são armazenados automaticamente os produtos já comprados, para assim o utilizador poder rapidamente encomendar novamente produtos que já adquiriu.

A página inicial é composta por uma lista das categorias de produtos, uma lista de lojas disponíveis para compras assim como uma lista dos produtos em saldo e os mais vendidos das respetivas lojas, sendo os produtos em saldo os de maior interesse por parte dos compradores.

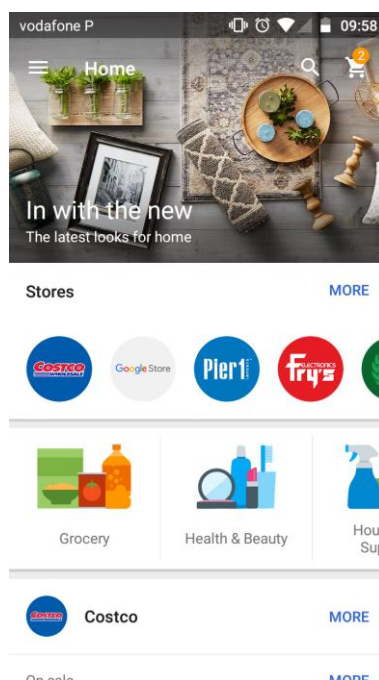


Figura 12. Página *Home* da aplicação Google Express.

Ao pesquisar por lojas assim como por categorias, é possível filtrar a pesquisa através da selecção de *tags* de categorias e subcategorias respetivamente. Em adição, também existem os filtros mais comuns (marca, preço, categoria, entre outros) e a opção de ordenar por vários critérios os resultados.

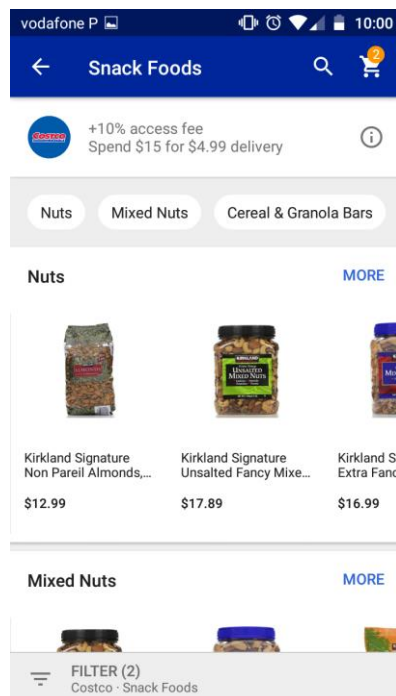


Figura 13. Filtragem de produtos da aplicação Google Express.

Para eliminar produtos do carrinho de compras é possível através do deslize do produto para o lado esquerdo ou diminuindo a quantidade do produto até zero .

Na pesquisa através de entrada de texto ou voz, os resultados que surgem são medianos de acordo com a semântica e a tentativa de pesquisar mais do que um produto em simultâneo não é possível com eficácia. Dependendo dos produtos a pesquisar, podem ou não surgir resultados de ambos os produtos mas bastante limitados, ou apenas de um produto.

### 2.4.2. Instacart

A Instacart nasceu em 2012 e é uma aplicação que permite ao utilizador comprar produtos de diferentes áreas. O utilizador consegue criar um pedido especial caso não consiga encontrar produto pretendido, criar receitas seleccionando diretamente produtos das lojas e adicionar receitas aos favoritos, assim como produtos.

Através da página inicial é possível pesquisar diretamente pela loja selecionada. Esta página apresenta, de acordo com a loja, uma secção de cupões de descontos, uma secção de recomendados para o utilizador e uma breve lista de produtos de cada categoria (Figura 14).

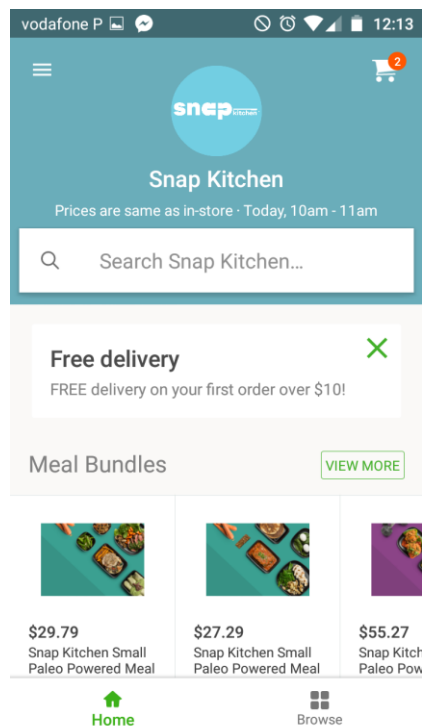


Figura 14. Página *Home* da aplicação Instacart.

Na pesquisa por categorias, surgem todas as subcategorias inseridas nesta, com uma breve lista de produtos. É possível filtrar a pesquisa através da seleção de *tags* das subcategorias. Por fim, ao entrar numa subcategoria, surgem todos os produtos dessa e é possível filtrar os resultados com filtros comuns e também ordená-los por diferentes critérios.

Um ponto interessante é o facto de os produtos adicionados ao carrinho de compras ficarem sinalizados e com a quantidade escolhida, podendo assim o utilizador lembrar-se que já adicionou o produto, durante a realização de compras.

Para eliminar produtos ou editar a quantidade do carrinho de compras é possível através do deslize do produto para o lado esquerdo ou abrindo mesmo. Aqui surgem duas opções para eliminar do carrinho e editar a quantidade do produto.

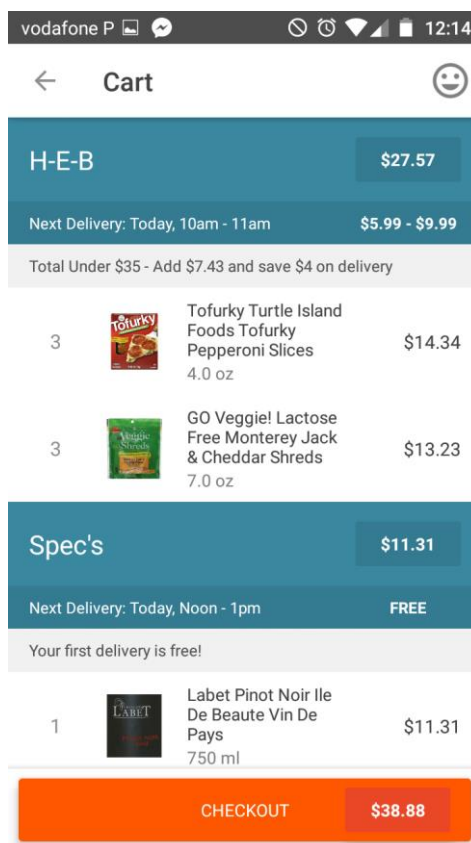


Figura 15. Carrinho de compras da aplicação Instacart.

Na pesquisa com entrada de texto ou voz, os resultados que surgem são medianos de acordo com a semântica e a tentativa de pesquisar vários produtos, em simultâneo, não apresenta resultados eficazes. Produtos que potencialmente não tenham qualquer relação, não devolvem resultados, sugerindo ao utilizador fazer um pedido especial. Nos produtos que possam estar relacionados, surgem resultados relacionados com todos os produtos mencionados, em simultâneo. Por exemplo com a pesquisa “chocolate e leite” surgem como resultados, leites com chocolate, chocolate de leite e outros que não está relacionado em nada do que o utilizador pretende.

### 2.4.3. FreshDirect

Esta aplicação está principalmente direccionada para compras de alimentação orgânica. A sua página inicial tem a opção de navegar diretamente na loja da aplicação, um menu para procurar produtos de forma rápida através das listas do utilizador e produtos já encomendados, assim como as recomendações da aplicação (Figura 16).

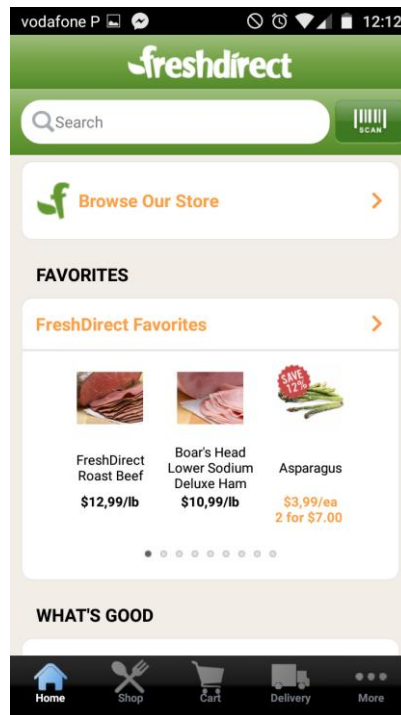


Figura 16. Página *Home* da plataforma FreshDirect

Na pesquisa por categorias, aparecem todas as subcategorias inseridas nesta, com uma breve lista de produtos. É possível filtrar a pesquisa através da seleção de *tags* das subcategorias. Por fim, ao entrar numa subcategoria surgem todos os produtos relativos a esta, e é possível filtrar os resultados com filtros comuns e também ordená-los por diferentes critérios.

Esta aplicação, assim como a Google Express, sinaliza os produtos que já estão adicionados ao carrinho de compras.

Para eliminar produtos do carrinho de compras, só é possível ao entrar diretamente na página na página do produto. Já para editar a quantidade, é possível através da página do produto ou através da lista no carrinho de compras (Figura 17).

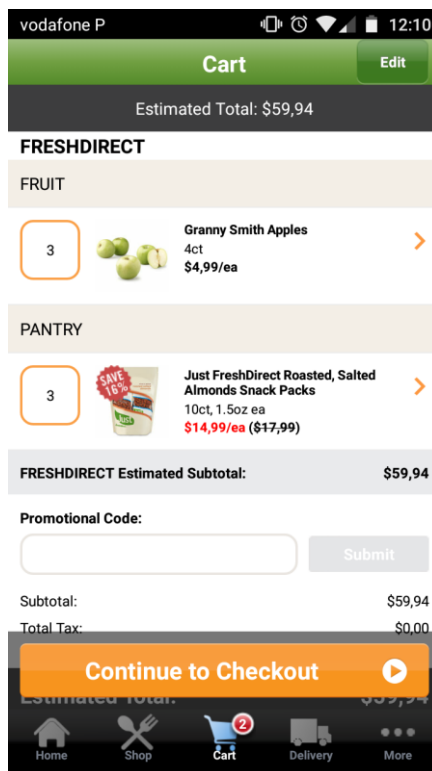


Figura 17. Carrinho de compras da plataforma FreshDirect.

Nesta aplicação a pesquisa por voz não está disponível. Assim, procurando os produtos por texto, os resultados apresentam uma coesão mediana de acordo com a semântica. Quanto à pesquisa relativa a mais do que um produto em simultâneo, acontece o mesmo que na aplicação Instacart, mencionada anteriormente. Em ambas surgem apenas resultados em certas combinações de produtos, e com estes integrados juntamente, não devolvendo resultados em separado para cada produto.

#### 2.4.4. Shipt

De modo ao utilizador conseguir efetuar compras de uma forma mais simples de produtos típicos de supermercado, surgiu em 2015 a aplicação Shipt. Na página inicial é possível procurar as lojas perto da localidade do utilizador, escolher a loja pretendida e pesquisar diretamente através da loja selecionada. Nesta página é apresentado um menu para pesquisar por categorias, comprar rapidamente através da secção de produtos já comprados e consultar os saldos do dia (Figura 18).

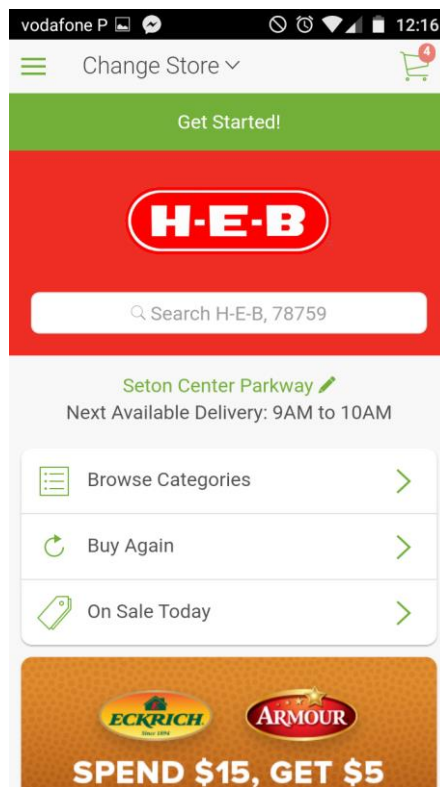


Figura 18. Página *Home* da aplicação Shipt.

Ao pesquisar por categorias, surge apenas uma lista com as subcategorias e só depois de seleccionar uma destas é que surgem os produtos da subcategoria. Nos resultados das subcategorias não é possível filtrar os resultados nem ordená-los. Os filtros comuns são possíveis através da pesquisa por voz ou texto, no entanto não há a opção de ordenar por critérios os resultados.

Esta aplicação também sinaliza os produtos adicionados ao carrinho de compras e a quantidade assim como na Google Express. No entanto, distingue-se das outras em relação à rapidez dos resultados, pois enquanto o utilizador está a escrever ou dizer o produto que deseja, os resultados vão surgindo em simultâneo.

Para alterar a quantidade ou eliminar produtos do carrinho de compras é necessário entrar diretamente na página do produto ou na própria lista do carrinho através da alteração da quantidade até zero.

Na pesquisa através de entrada de voz ou texto, os resultados que surgem são medianos de acordo com a semântica e apresentam também uma eficácia mediana na tentativa de pesquisa de vários produtos em simultâneo (Figura 19).



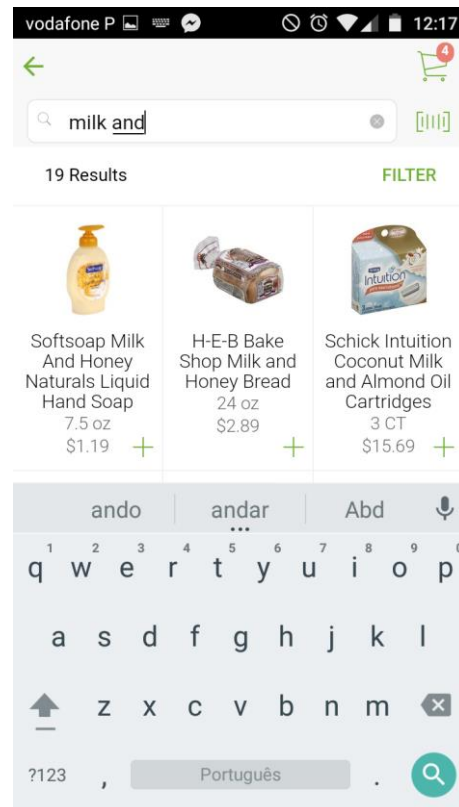


Figura 19. Resultados surgidos durante a entrada do utilizador

Os resultados estão dependentes dos produtos pesquisados, pois pode surgir uma lista com os produtos em separado. No entanto também pode surgir uma lista com produtos não relacionados e produtos que sejam a fusão de ambos, como por exemplo iogurtes de banana, quando a pesquisa é “iogurtes e bananas”. A aplicação também dá a opção ao utilizador de fazer um pedido especial, caso não encontre o que deseja.

### 2.4.5. Síntese da Análise

Tendo como base os critérios avaliados mais importantes em relação à Shelf AI, foi criada uma tabela onde é possível verificar esses critérios, para fazer uma breve conclusão da análise realizada previamente.

Tabela 1. Comparação de acordo com alguns critérios das aplicações analisadas.

Funcionalidades e características	Google Express	Instacart	FreshDirect	Shipt	Shelf AI
Pesquisa por voz	✓	✓	✗	✓	✓
Scan Código de Barras	✗	✗	✓	✓	✓
Pesquisa de vários produtos	○	✗	✗	○	✓
Resultado coeso com semântica	○	○	○	○	✓
Filtros de pesquisa	✓	✓	✓	✓	✓
Rapidez de resposta	✓	○	○	✓	✓

✓ sim   ○ parcial   ✗ não

## 2.5. Síntese do Capítulo

Neste capítulo foi apresentada a revisão da literatura relacionada com interfaces conversacionais.

Uma interface conversacional é baseada em inteligência artificial, capaz de interagir com humanos através de linguagem natural por voz (assistentes de voz) ou texto (*chatbots*).

O trabalho de Alan Turing - A máquina de Turing e O jogo da Imitação - potenciou a criação de interfaces inteligentes com capacidade de simularem humanos. Desta forma, foi criado o concurso Loebner Prize em que o seu objetivo passa por eleger o *chatbot* com melhor capacidade de simular um humano no diálogo. No decorrer dos anos, surgiram *chatbots* que se distinguiram e marcaram a evolução destas interfaces: ELIZA, o primeiro *chatbot* e vencedor do Loebner Prize,

PARRY, em resposta à criação de ELIZA e A.L.I.C.E., um *chatbot* mais recente e também vencedor do Loebner Prize.

Atualmente é visível que as interfaces conversacionais sofreram uma grande evolução.

Os *chatbots*, que interagem através de texto, são cada vez mais comuns e fáceis de aceder pois estão maioritariamente disponíveis em plataformas de chat como Facebook Messenger, Kik e WeChat. Tay e Zo, dois *chatbots* desenvolvidos pela Microsoft foram criados com o objetivo de desenvolverem as suas capacidades de fala, à medida que interagissem com os humanos.

Os assistentes de voz como a Cortana, Alexa e Google Assistant, comunicam através de voz e funcionam maioritariamente como assistentes pessoais. São capazes de realizar diversas tarefas, como definir lembretes, reproduzir música, apresentar meteorologia, realizar chamadas, entre outras.

Outro tipo de interface conversacional são as interfaces Human-Powered. Dois exemplos destas são a Magic e a Operator, que permitem ao utilizador interagir com agentes humanos apoiados pela Inteligência Artificial, para responderem aos pedidos. A Inteligência Artificial tem como finalidade personalizar a interação, assim como auxiliar o agente humano a gerar respostas mais rápidas. Isto é possível, pois é armazenada toda a informação importante do utilizador, como por exemplo os seus produtos preferidos.

As compras *online* podem ser as preferidas do consumidor devido à sua maior facilidade de compra, assim como rapidez. No entanto, mesmo este método pode-se tornar mais longo. Em contrapartida, as empresas estão a apostar na criação de interfaces de comércio conversacional, a fim de ficarem mais próximos dos seus clientes, assim como melhorar o processo de compra destes. Estas plataformas podem ser *bots*, agentes humanos, serviços ou agentes humanos com apoio à inteligência artificial.

De acordo com os dados do BI Intelligence (2016), o número de utilizadores de aplicações de mensagens está a aumentar, comparativamente às redes sociais.

A IHC é importante para a criação de uma interface agradável para o utilizador, pois foca-se na forma como ocorre a interação entre o humano e a máquina, para providenciar uma boa experiência ao utilizador durante a interação.

A usabilidade associada à IHC está relacionada com a facilidade com que o utilizador se depara ao interagir com uma interface. Antes de uma interface ser implementada, o mais apropriado é testar a usabilidade desta. Existem vários métodos de teste, mas Nielsen (2012) considera que o teste mais eficaz é o teste ao utilizador.

As interfaces conversacionais, comparativamente às interfaces gráficas são mais complexas em termos de criação, uma vez que apresentam pouco ou nenhum conteúdo visual. Consequentemente, é necessário recorrer a outras técnicas para melhorar a interação. O *designer* é responsável por determinar as *prompts*, gramáticas e ações e deve ter em conta fatores humanos e de ambiente, que podem afetar a usabilidade e a forma de interação.

## 3. Análise e Design

O objetivo do projeto da presente dissertação passa por transformar a plataforma Shelf AI numa interface conversacional híbrida, onde o utilizador poderá optar por interagir por voz e/ou texto/toque, assim como a interface responder através de conteúdos visuais e por voz. Em consequência, para a transformação da aplicação atual, foi inicialmente realizado um questionário de forma a compreender o conhecimento dos potenciais utilizadores, tal como a sua opinião relativamente a interfaces conversacionais para compras, bem como a definição do público-alvo para deste modo servir como fundamento para outros pontos da aplicação. Posteriormente foram definidos os fatores diferenciadores em relação às outras aplicações de compras que estão interligados com as funcionalidades da Shelf AI, onde são apresentadas funcionalidades do ponto de vista do utilizador, assim como do ponto de vista do sistema.

Após a definição destes elementos, foi iniciada a fase do *design* de usabilidade conversacional onde são abordadas as alterações a serem realizadas na Shelf AI, de forma a transformá-la numa interface conversacional híbrida a mais otimizada possível. O ponto fulcral neste tipo de *design* é o fluxo de diálogo, onde foram definidas as *prompts* do sistema de acordo com as entradas do utilizador, para saber como e quando a interface comunica com o utilizador dependentemente dos contextos.

### 3.1. Questionário Inicial

Antes de serem definidos os pontos relacionados com a Shelf AI e assim passar à fase de *design* e realização de protótipos, surgiu a necessidade de realizar um questionário com o intuito de conhecer os possíveis utilizadores e tomar conhecimento da sua opinião relativamente a uma interface conversacional para compras.

Os pontos mais importantes passaram por compreender o nível de conhecimento que os participantes têm em relação às interfaces, o seu nível de interesse em recorrer a estas para realizarem compras, assim como quais as vantagens e utilidades destas para eles.

O questionário foi criado através do *Google Forms* e disponibilizado *online* através do *e-mail* e redes sociais. Este contou com um total de 233 participantes e era composto por 11 questões quantitativas e qualitativas a fim de obter respostas fechadas e abertas, respetivamente.

Algumas das questões tinham como objetivo a obtenção de dados demográficos dos participantes como a idade, género, profissão e nacionalidade. Estas questões levaram a concluir que a maioria dos participantes (71,21% - 166 participantes) têm idade entre 18 a 24 anos, tal como se pode consultar na Figura 20 e 54,1% (126 participantes) são do sexo feminino.

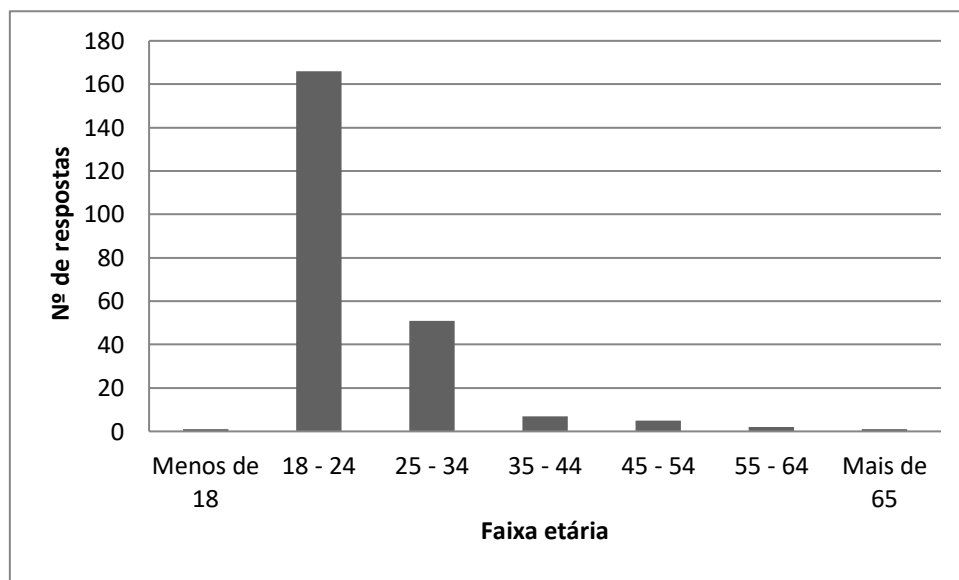


Figura 20. Representação das faixas etárias dos participantes do questionário

Em relação às nacionalidades, 92,7% (216) são de nacionalidade portuguesa enquanto apenas 7,3% (17) são de outras nacionalidades: inglesa, brasileira, britânica, belga, francesa, checa, romena, búlgara, eslovaca e argelina.

Quanto à profissão dos participantes, também foi encontrada uma grande variedade em profissões: estudantes, programadores, engenheiros, *designers*, contabilistas, docentes, técnicos de informática, investigadores, geógrafos, entre outros.

Como é possível conferir, os participantes apresentavam uma grande variedade no seu perfil em relação à idade, nacionalidade e profissão. Estes elementos de diferenciação, permitiram assim obter opiniões diversificadas nas questões, de escolha múltipla e quantitativas, relacionadas com as interfaces conversacionais, que são abordadas a seguir.

Tabela 2. Questões de escolha múltipla e quantitativas

Questões
1. Quais dos seguintes tipos de interfaces conversacionais já utilizou?
2. Qual considera a melhor forma de o utilizador interagir com uma interface conversacional?
3. Qual considera a melhor forma para a interface conversacional comunicar com o utilizador?
4. Qual o nível de interesse em realizar compras <i>online</i> através de um diálogo numa interface conversacional?
5. Qual o nível de utilidade que teria uma interface conversacional para realizar compras <i>online</i> ?

A primeira questão, relacionada diretamente com as interfaces conversacionais, passou por compreender se os participantes já tinham utilizado alguma interface conversacional com o objetivo de perceber se as respostas seguintes, relativamente a opiniões pessoais sobre estas interfaces, eram fundamentadas em experiências ou não (Figura 20). Com esta questão, concluiu-se que 32,6% (76) dos participantes nunca utilizaram nenhum tipo de interface conversacional, enquanto que 33% (77) já utilizaram *chatbots* e em minoria encontra-se a interface de voz que foi utilizada apenas por 8,2% (19) dos participantes.

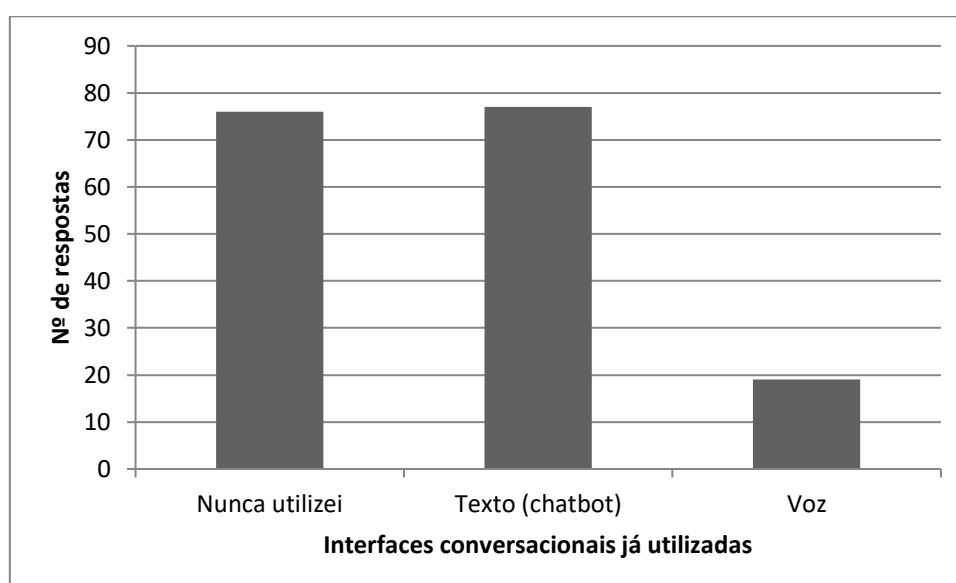


Figura 21. Número de utilizações de interfaces conversacionais

O questionário era também composto outras duas questões, de escolha múltipla, com o intuito de tomar conhecimento das preferências do utilizador em relação à interação com que se deparam ao utilizar uma interface conversacional.

Um das questões abordava o modo com que o utilizador prefere interagir com estas interfaces, ou seja, por texto/toque, voz ou de forma híbrida, isto é através de texto/toque e voz (Figura 22). Quase metade dos participantes – 109 participantes (46,8%) – preferem interagir de forma híbrida, enquanto que 93 participantes (39,9%), uma grande parte, prefere interação por texto/toque e apenas 31 participantes (13,3%) têm preferência pela interação por voz.

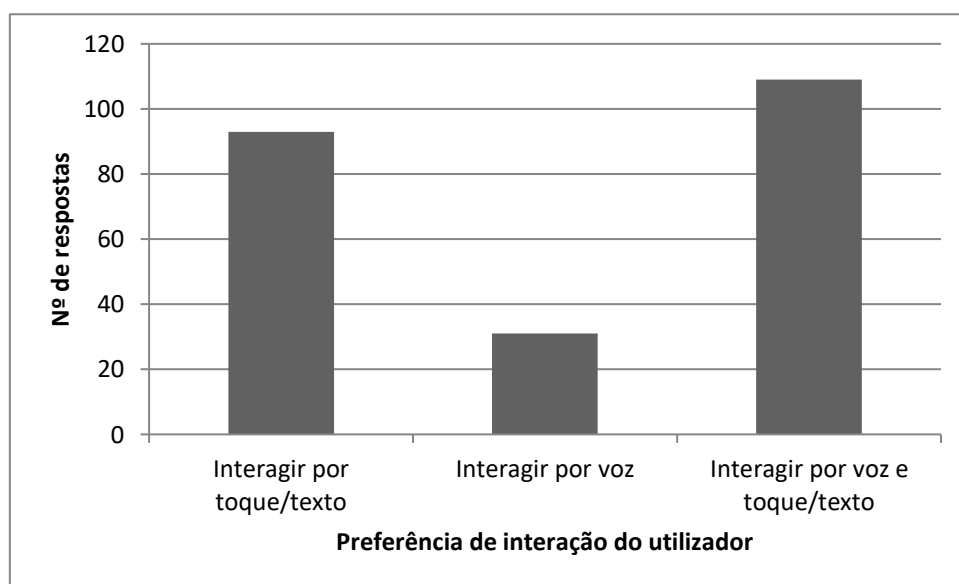


Figura 22. Preferência do utilizador de como interagir com a interface

Na questão seguinte, relativamente à preferência dos utilizadores quanto à forma como a interface comunica com estes, mostrou ser semelhante à conclusão da questão anterior (Figura 23). A maioria dos participantes (91,4% – 213) – preferem que a interface comunique de forma híbrida, ou seja, por voz e texto/elementos visuais, enquanto que 6,9% (16) dos participantes optam pela interface que comunique por texto/elementos visuais e por sua vez, apenas 1,7% (4) participantes têm preferência pela comunicação por voz.

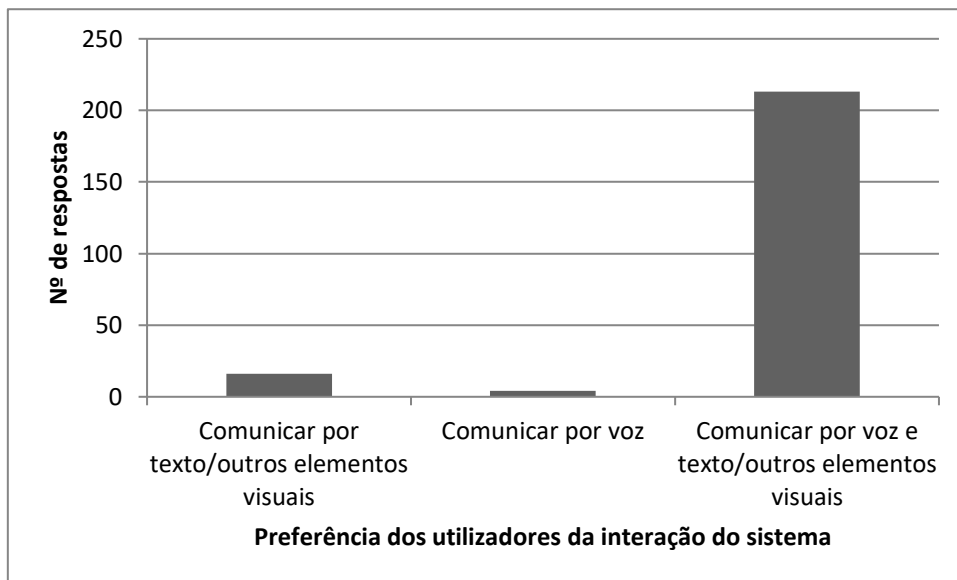


Figura 23. Preferência dos utilizadores do modo de interagir do sistema

De seguida, também foram realizadas duas questões de caráter quantitativo com o objetivo de compreender o nível de interesse e de utilidade que os participantes têm nas interfaces conversacionais.

Em relação ao nível de interesse para realizar compras através do diálogo numa interface conversacional híbrida, 32,6% (76) dos participantes apresentaram um interesse mediano, de seguida 25,3% (59) mostraram um interesse elevado, enquanto que 21% (49) participantes têm baixo interesse e 18,5% (43 participantes) não têm interesse. Num total de 233 participantes, apenas 2,6%, que corresponde a 6 participantes, é que demonstram um interesse máximo em realizar compras desta forma, no entanto, foram mais participantes a responderem que têm médio interesse do que baixo interesse.



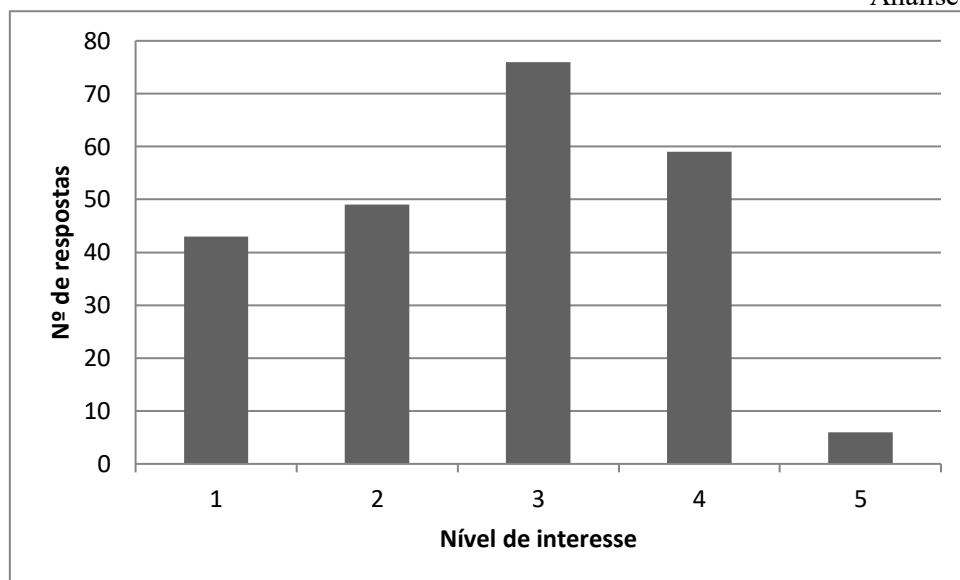


Figura 24. Nível de interesse dos utilizadores quanto a compras através de um diálogo

Apesar do interesse dos participantes ser mediano, tal como mencionado anteriormente, notou-se quanto à opinião dos utilizadores em relação a estas interfaces para realizar compras, que o nível utilidade destas é o mais elevado. Foi realizada uma análise individual de forma a comparar as respostas da questão anterior com esta e foi notório que muitos dos participantes que não têm interesse elevado nestas interfaces, reconhecem que estas podem ser úteis em contexto de compras.

Dos participantes, 36,8% (86) consideram estas interfaces úteis, 23,2% (54) consideram que têm utilidade mediana, enquanto que 16,7% (39) e 14,6% (34) dos participantes consideram a interface com pouca ou nenhuma utilidade, respetivamente. Apenas 8,6% (20) consideram este tipo de interface muito útil mas como já referido, notou-se uma melhoria nas respostas da utilidade em relação ao interesse, onde as respostas positivas em relação a estas interfaces continuaram a mostrar favoritismo, tal como na anterior (Figura 25).

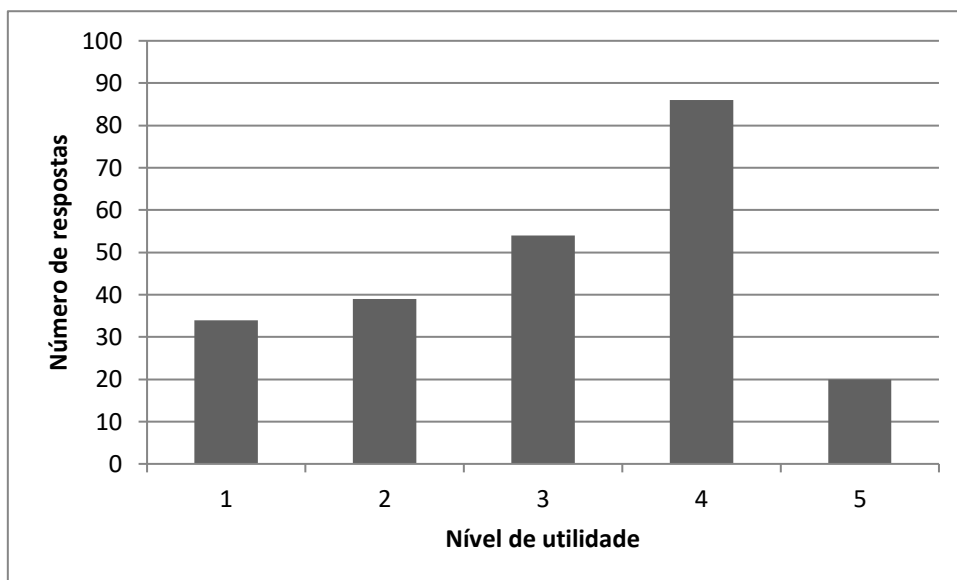


Figura 25. Nível de utilidade de uma interface conversacional quanto a compras

Por fim, foi apresentada uma questão de resposta aberta, que passava por obter a opinião dos utilizadores, para compreender quais as vantagens das interfaces conversacionais, que estes consideram de forma a terem utilidade. Alguns participantes mencionaram os motivos pelas quais estas interfaces podem ser úteis, tal como está representado na Tabela 3.

Tabela 3. Vantagens associadas às interfaces híbridas para compras

Vantagens
1. Ajuda no esclarecimento de eventuais dúvidas rapidamente.
2. Facilita o processo de compra para pessoas com limitações físicas e invisuais.
3. Experiência de compra mais intuitiva para os menos aptos em tecnologia pois a experiência entre o homem-máquina torna-se humanizada e natural.
4. Simplifica e acelera o processo de compra repetitivo para o utilizador.
5. Facilita o processo de pesquisa do produto desejado.
6. Atendimento e resultado personalizado consoante as necessidades, hábitos e exigências do utilizador.
7. Maior proximidade entre o utilizador e o retalhista.
8. Substitui a troca de <i>emails</i> para obter informações dos produtos e outras informações relacionadas com a compra.

Esta questão de resposta aberta, permitiu compreender o porquê de os potenciais utilizadores considerarem útil uma interface conversacional para efetuarem compras. Em termos de projeto, esta conclusão veio corroborar a utilidade que as interfaces híbridas possuem.

### 3.2. Público-alvo

Um estudo efetuado pelo Instituto Nacional de Estatística (INE) (Instituto Nacional de Estatística, 2016), em relação à quantidade de indivíduos com idades compreendidas entre os 16 e 74 anos que já realizaram compras *online* para fins privados, mostra que no primeiro trimestre de 2016 os indivíduos que recorreram mais ao *e-commerce*, têm idades compreendidas entre os 16 e 54 anos.

Relacionando os resultados do estudo do INE com os obtidos no questionário inicial, apresentado no subcapítulo anterior, foi idealizado o público-alvo da Shelf AI. O público-alvo passa pelos utilizadores com faixa etária compreendida entre os 18 e 55 anos, sendo que neste grupo existe uma maior probabilidade em possuírem interesse em realizar compras *online*. A

possibilidade de haver uma interação por voz nesta aplicação, faz com que os utilizadores menos experientes no uso de *smartphones*, possam aceder à aplicação de uma forma mais facilitada.

É notório que o público com maior interesse é aquele que necessita realizar compras de uma forma mais rápida e fácil, podendo poupar tempo e executar mais do que uma tarefa em simultâneo.

### 3.3. Fatores Diferenciadores

A Shelf AI, como interface conversacional híbrida, consegue diferenciar-se pelas suas características e funcionalidades, quando comparada com outras interfaces destinadas ao *e-commerce*. As interfaces gráficas analisadas no subcapítulo 2.4, não proporcionam ao utilizador uma interação híbrida na qual é possível a interação através de um diálogo natural, “*hands-free*”<sup>7</sup> para realizar mais do que uma tarefa em simultâneo, assim como a possibilidade de retirar dúvidas diretas interagindo com a interface.

Quanto à assistente de voz, por exemplo a Alexa que está integrada em dispositivos como Amazon Echo, Echo Dot, e Amazon Tap, em comparação às aplicações abordadas no subcapítulo 2.4, é composta por inteligência artificial, permitindo ao utilizador interagir através de linguagem natural e de forma personalizada. Uma das suas principais finalidades passa por permitir ao utilizador a realização de compras na *Amazon*, no entanto, apresenta limitações associadas à sua interação. Essas limitações estão relacionadas com o facto de esta ser uma interface que interage apenas por voz nos dispositivos em que se insere. Esta forma única de interação pode não ser a mais adequada para determinados momentos e contextos. Por exemplo, a sua utilização num ambiente ruidoso poderia tornar-se num problema para a interação e experiência do utilizador durante o processo de compras.

Consequentemente, a Shelf AI sendo uma interface híbrida baseada em inteligência artificial, proporciona ao utilizador mais do que uma forma de realizar compras, tendo este o poder de escolher a forma que prefere para efetuar as compras através da pesquisa de produtos por voz ou por texto. Permite também conferir os resultados de forma igual, obtê-los consoante as exigências, necessidades e perfil do utilizador, retirar dúvidas de forma nítida e rápida, assim como obter sugestões. Isto faz com que a Shelf AI se destinga das outras interfaces em relação a determinados pontos apresentados anteriormente.

---

<sup>7</sup> Termo inglês utilizado para definir que o utilizador mantém as mãos livres.

### 3.4. Funcionalidades

Após uma análise do estado de arte em relação às interfaces conversacionais existentes e às plataformas que permitem efetuar compras, tal como é o caso das aplicações gráficas e dispositivos de voz – Amazon Echo (Alexa) e Google Home (Google Assistant) - foram determinadas funcionalidades para a Shelf AI, de forma a distinguir-se dos concorrentes. Tendo em conta as opiniões e necessidades dos utilizadores, foi possível confirmar que a maioria das funcionalidades previamente definidas iam de encontro aos interesses do utilizador, não sendo assim necessário grandes reformulações das funcionalidades.

Na Tabela 4 são apresentadas as funcionalidades que o utilizador pode usufruir diretamente ao interagir com a interface.

Tabela 4. Funcionalidades do ponto de vista do utilizador

Funcionalidades para o utilizador
<i>Login/logout</i> na aplicação.
Pesquisar produtos através de voz, texto/toque ou código de barras.
Pesquisar produtos recorrendo a características/filtros.
Alterar quantidade do produto antes de adicionar ao carrinho de compras.
Tirar dúvidas sobre produtos através de voz.
Fazer <i>checkout</i> através de voz ou texto.
Obter informações sobre encomenda através de voz ou acedendo externamente ao email.
Opção para ativar e desativar comunicação da interface por voz.
Vizualização de <i>pop ups</i> do reconhecimento de voz.
Criação de lista de compras.
Alterar dados da conta.
Adicionar produtos ao carrinho de compras.
Eliminar produtos do carrinho de compras.
Alterar quantidade do produto no carrinho de compras.
Requisitar apresentação de resultados por voz.

Estas funcionalidades baseiam-se essencialmente na definição das ações e distinguem-se na forma de realizá-las, ou seja, definindo se podem ser realizadas por voz, toque/texto ou de forma híbrida.

Em relação às funcionalidades do sistema para com o utilizador, tal como estão representadas na Tabela 5, são ações que o sistema toma por iniciativa própria.

Tabela 5. Funcionalidades do ponto de vista do sistema

Funcionalidades para o sistema
Pedir informações ao utilizador por voz e graficamente.
Apresentar produto de topo por voz ao utilizador.
Apresentar resultados de todos os produtos graficamente.
Apresentar sugestões por voz ao utilizador.
Apresentar resultados personalizados consoante o perfil do utilizador.
Iniciativa de auxiliar o utilizador na filtragem de resultados.

A iniciativa do sistema permite que o utilizador não tenha que iniciar sempre a interação, permitindo assim que este usufrua de uma melhor experiência.

### 3.5. Fluxos de Interação

Após a determinação das funcionalidades do sistema do ponto de vista do utilizador assim como do ponto de vista do próprio sistema, foram realizados diagramas UML, de forma a representar as diferentes saídas que as ações dos utilizadores têm, alterando assim a reação do sistema de acordo com essas.

Para realizar o *login* na aplicação, é necessário o utilizador inserir todos os dados necessários para iniciar a sessão com sucesso. Enquanto não inserir, não é possível avançar para o início da sessão. Após o início da sessão, o utilizador consegue entrar na página inicial da aplicação onde pode ter acesso a todas as suas funcionalidades, o que não é possível se não realizar o *login* (Figura 26).

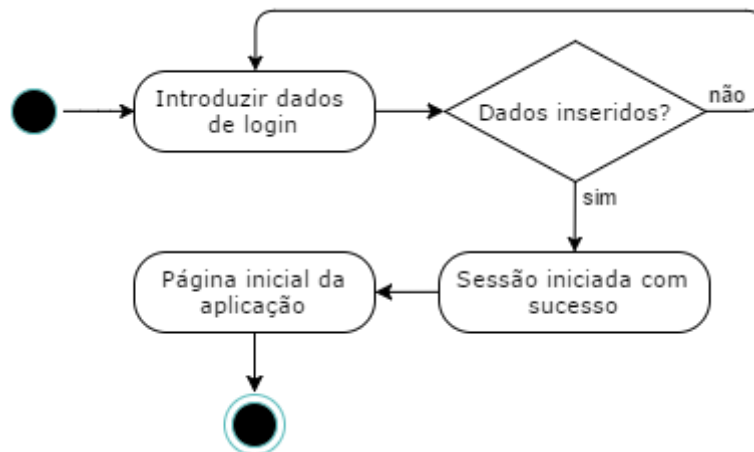


Figura 26. Início de sessões na Shelf AI

Caso o utilizador necessite ou apenas deseje desligar a comunicação por voz do assistente, ou seja, não ouvir as *prompts* durante a sua interação, pode recorrer às definições para ativar ou desativar a opção da comunicação por voz, sempre que quiser (Figura 27).

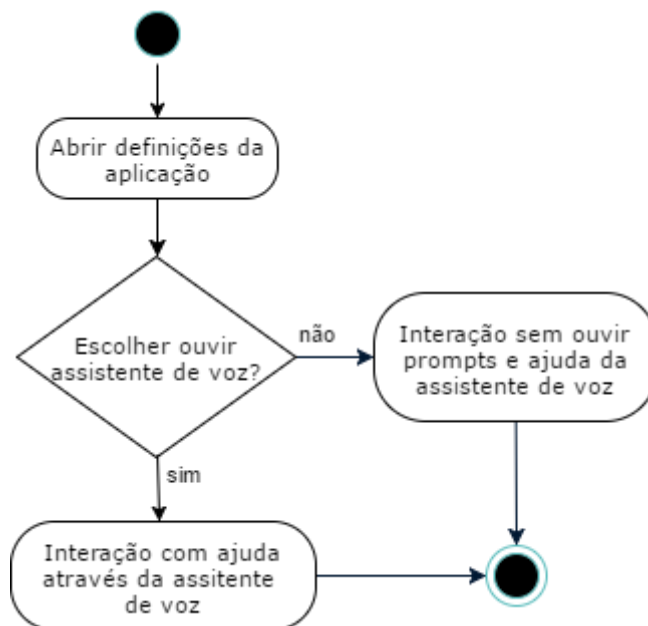


Figura 27. Ativar ou desativar de interação por voz

Após o *login*, uma das funcionalidades que o utilizador tem ao seu dispor é a pesquisa de produtos. Para prosseguir para esta funcionalidade, é necessário que o utilizador insira o nome do produto que pretende procurar. Caso a entrada seja inválida, surgirá um *pop up* informando que ocorreu um erro na entrada do utilizador. Independentemente da entrada ser por voz ou texto, o erro poderá ser corrigido por uma das formas de entrada. Esse erro pode estar relacionado com dificuldades no reconhecimento de voz, no caso da entrada ser realizada desta forma ou erro de escrita no caso da entrada ser realizada por texto.



Uma vez validada a entrada, será retornada a lista inicial de produtos disponíveis. Se a pesquisa necessitar de ser filtrada, no caso de existirem muitos produtos de acordo com o que o utilizador pretende, este terá que dar uma nova entrada de forma a que seja adicionada alguma característica do produto que deseja, como por exemplo a marca. Após o primeiro filtro, será apresentada uma lista de produtos mais reduzida, no entanto, se esta continuar extensa será recomendado filtrar novamente a pesquisa, visto que o objetivo é que o utilizador chegue a um resultado restrito, conseguindo assim obter o produto desejado. Ou seja, caso ocorra uma nova filtragem, retornará outra lista de produtos. Se não for necessário recorrer a mais filtros, termina assim a pesquisa e é possível aceder à lista final (Figura 28).

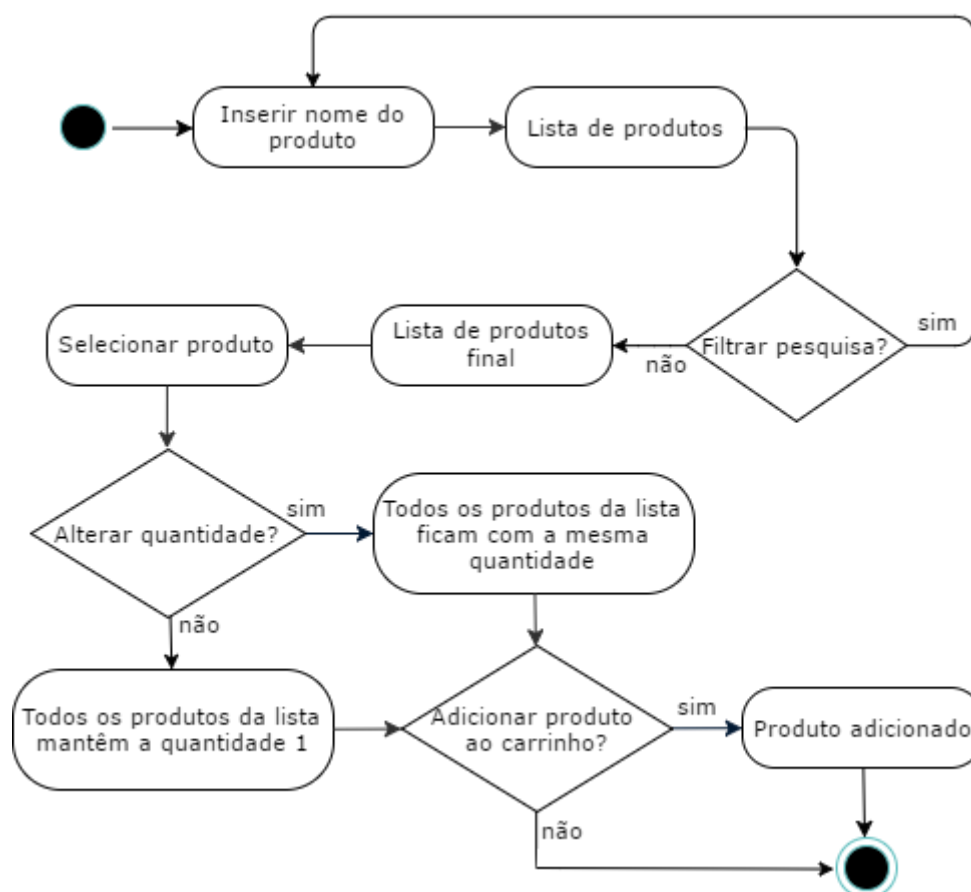


Figura 28. Pesquisa de produtos independentemente através de voz ou texto

Durante a pesquisa de produtos, o utilizador tem a opção de selecionar a quantidade desejada e desta forma simular o preço resultante da escolha após essa, adicionar ao carrinho já com a quantidade que deseja. Caso o utilizador não altere, o valor por defeito da aplicação é de uma unidade.

Para realizar o *checkout* (Figura 29) é necessário o utilizador passar por três etapas de introdução de dados:

1. Introduzir dados pessoais (nome, número, e-mail);
2. Inserir os dados de entrega (morada, dia e hora);
3. Escolher método de pagamento e adicionar os dados (número de cartão, código de segurança, validade).

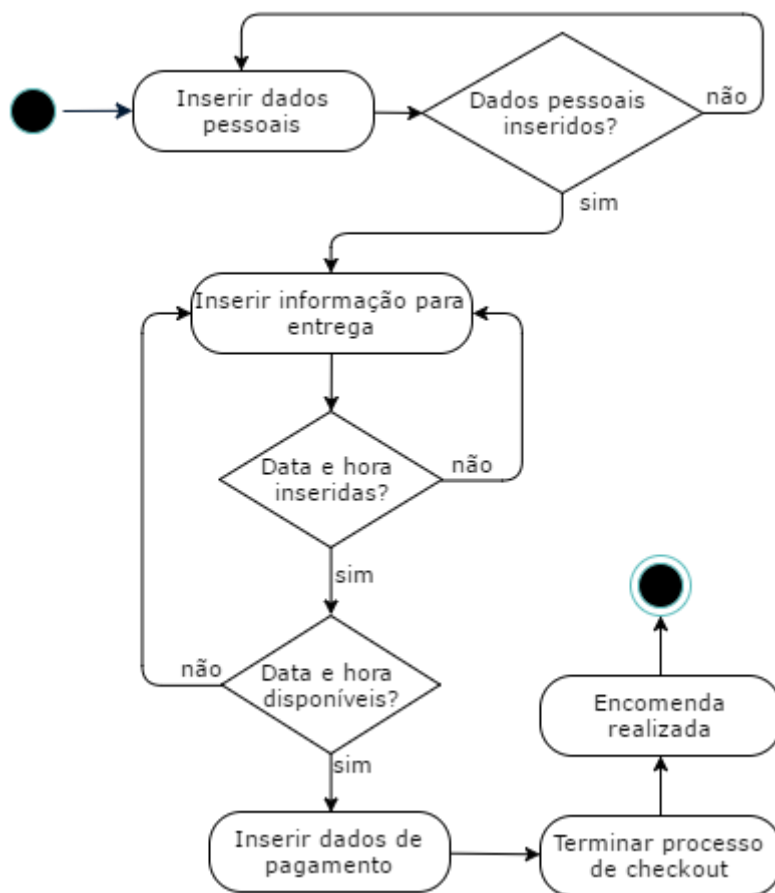


Figura 29. Processo de *checkout*

Neste sentido, sempre que alguma entrada de dados não for válida, os campos com erro ficam a vermelho e é necessário o utilizador repetir o processo. Se a entrada inicial tiver sido por voz ou texto, o utilizador poderá fazer a correção recorrendo novamente à voz ou caixa de texto. Se por sua vez na entrada por texto ou voz for inserido algum dado inválido, o utilizador poderá repetir o processo com recurso à forma híbrida.

### 3.6. Design de Usabilidade Conversacional

Após a análise do questionário, assim como a definição dos pontos diretamente relacionados com a Shelf AI – público-alvo, fatores diferenciadores e funcionalidades – foi possível passar à

fase do *design*. A Shelf AI foi alvo de constantes avaliações relacionadas com a interação, com o intuito de compreender qual o comportamento que esta poderia ter de forma a tornar-se numa interface conversacional híbrida.

Posto isto, foram identificados os pontos a alterar e a acrescentar na Shelf AI, sendo que estes foram sujeitos a várias mudanças durante todo o processo. Por conseguinte, foi possível compreender, de uma forma mais prática e concreta, aquilo que envolve a criação de uma interface conversacional híbrida e como é possível tirar partido deste tipo de interface.

As interações feitas somente por voz, tanto em entradas como saídas, não são se adequam a todas as tarefas (Dybkjær & Bernsen, 2001) e momentos, por isso uma boa solução é oferecer ao utilizador a possibilidade de interagir por voz e/ou por toque/texto. Tal como o utilizador, o sistema deve transmitir informação e interagir de forma híbrida, ou seja, através da voz e conteúdos gráficos/texto, para se adequar a diferentes contextos de utilização.

Segundo Nielsen (2012), a usabilidade de um sistema está relacionada com a facilidade com que o utilizador se encontra ao interagir com uma interface. No entanto, as necessidades do público-alvo são imprescindíveis no conceito de usabilidade, pois caso a interface não seja útil para satisfazer tais necessidades, a usabilidade deixa de ter grande relevância. O *design* da Shelf AI foi desenvolvido com base neste conceito de usabilidade, sendo que para a definição da aplicação foram tidas em consideração as respostas dos participantes do questionário inicial.

### 3.6.1. Alterações na Aplicação

Para converter a aplicação Shelf AI numa interface conversacional híbrida foram necessárias várias alterações. Essas alterações consistem principalmente no *design* de fluxo de diálogo, para que seja possível interagir de forma natural com a interface. A funcionalidade de interagir por voz levou a outras modificações.

O sistema deve fornecer um *feedback* ao utilizador durante as ações como reconhecimento de voz e erros ocorridos, pois caso o utilizador proceda na interação com receio e dúvidas, pode influenciar a sua interação e satisfação. Um dos *feedbacks* importantes está relacionado com o reconhecimento da voz, para que o utilizador tenha conhecimento de que o sistema compreendeu a sua entrada, podendo assim prosseguir para a próxima ação e sem qualquer dúvida relacionada com o reconhecimento do sistema.

O *feedback* que existe atualmente na Shelf AI surge quando o utilizador recorre a um *botão* para adicionar produtos. Automaticamente, esse *botão* ativa o reconhecimento de voz esperando a entrada do utilizador, onde em simultâneo surge na caixa a indicação “a ouvir...”. Após a entrada do utilizador, aparece o resultado do *Text to Speech* (TTS) de acordo com o que foi dito (Figura 30). Assim, o utilizador pode confirmar se ocorreu uma compreensão eficaz do sistema, assim como alterar manualmente ou por voz através da edição da entrada na caixa de texto ou utilizado o *botão* para edição da entrada por voz, repetivamente.



Figura 30. *Feedback* da Shelf AI original no reconhecimento de voz

Apesar de a aplicação já conter este *feedback*, surgiu a necessidade de alterar o modo como este é dado. Foi criado um *pop up* para que o utilizador possa ter a certeza de que o sistema está a escutá-lo, bem como a reconhecer a entrada por si inserida. Esse *pop up* surge quando o utilizador faz uso do comando “Shelf” ou clicando no *botão* destinado para a entrada de voz. O *pop up* apresenta inicialmente um ícone relacionado com ondas de som, para o utilizador compreender que a Shelf AI está pronta escutá-lo. Quando este começa por ditar a sua entrada, o ícone é animado consoante o seu tom de voz. Com essa mesma animação do ícone, surgem, em simultâneo e de forma sequencial palavra a palavra, para assim o utilizador obter o *feedback* do reconhecimento da sua entrada em tempo real (Figura 31).

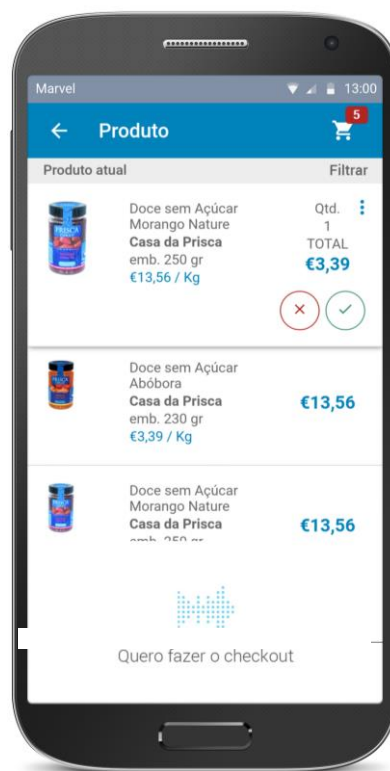


Figura 31. Alternativa ao *feedback* através de *pop ups*

Os sistemas com iniciativa própria são ideais para quando necessitam de partes de informação, enquanto que os sistemas com iniciativa do utilizador são mais complicados para o utilizador que não tem experiência (Dybckjær & Bernsen, 2001). Deste modo, foi definido que a Shelf AI será um sistema de iniciativa mista, uma vez que o sistema tomará a iniciativa de pedir informação ou guiar o utilizador, como por exemplo através da oferta de sugestões, sem serem requisitadas pelo utilizador, assim como fazer questões para ajudar no processo de pesquisa principalmente para os utilizadores não experientes. Quanto ao utilizador, quando este é experiente, pode também tomar iniciativa caso precise de informações ou de corrigir algum erro.

Atualmente na Shelf AI só é possível realizar o *checkout* acedendo ao *site* do retalhista, surgindo apenas um *pop up* na aplicação, informando que para realizar o *checkout* é necessário aceder ao *site* do retalhista utilizando os dados da sua conta (Figura 32).

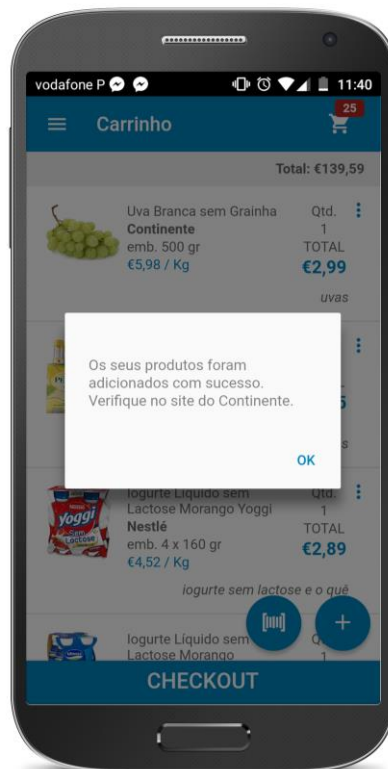


Figura 32. *Pop up* do *checkout* original da Shelf AI

Para realizar o *checkout* na versão híbrida, foi decidido incluir este processo na própria aplicação. Aceder externamente ao *site* conduz a que o processo seja mais demorado, cria quebras na interação e impossibilita o auxílio proveniente do sistema para ajudar na marcação de entrega. Assim sendo, foi criado um ecrã de forma a materializar o *checkout* graficamente, onde estão situados os campos necessários a preencher para a realização do mesmo, tal como é possível consultar um dos exemplos na Figura 33.

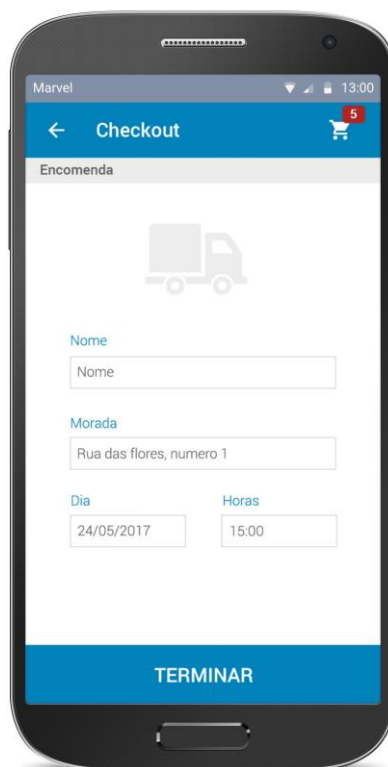


Figura 33. Ecrã do processo de *checkout*

Neste processo são apresentadas as *prompts* para pedir ao utilizador os dados necessários para esses mesmos campos, onde este tem a possibilidade de responder por voz ou por texto. Para além das *prompts* de pedido de informação, também podem surgir outras para caso exista algum erro de introdução de dados, ou caso não seja possível fazer marcação da entrega para uma determinada data e hora.

### 3.6.2. Fluxos de diálogo

O diálogo, como já referido anteriormente, é uma das principais novidades em comparação à aplicação existente, exigindo assim especial atenção e dedicação ao seu *design*.

Tendo em conta as diferentes funcionalidades e para compreender como pode ocorrer o diálogo, foram desenhados fluxos através de diagramas, abordando assim o *login*, troca de dados, dúvidas, pesquisa de produtos e resultados, carrinho de compras, *checkout* e pedido de informações.

Durante a definição dos fluxos do diálogo, foram também realizados alguns testes na ferramenta Api.ai. Estes testes serão abordados no subcapítulo 4.1, uma vez que fazem parte do processo de realização dos protótipos.

Para a criação dos fluxos, foi fundamental ter em consideração como o utilizador realiza compras e qual o tipo de diálogo tomado por este, tendo em conta que pode interagir através de

linguagem natural. Foram também pensadas entre duas a três alternativas para as *prompts*, dependendo da funcionalidade em causa, de forma a que estas apresentem uma variedade no discurso, evitando que se torne numa interação repetitiva e assim, que o utilizador não se aborreça. Estas mesmas *prompts* sofreram uma grande evolução ao longo do processo do *design*, pois durante os testes de diálogo, realizados na ferramenta Api.ai, começaram por ser visíveis certos problemas relacionados com a repetição. Com isto, foi possível perceber que uma boa solução era personalizar as *prompts* de filtragem ao máximo, de acordo com cada produto, de forma a existir uma maior variedade e não se tornar uma experiência monótona e exaustiva.

Inicialmente foram escolhidos certos produtos para desenhar o diálogo do sistema para a pesquisa e filtragem, determinando assim que os principais filtros para a maioria dos produtos seriam o tipo e a marca, onde é possível consultar na Tabela 6.

Tabela 6. Filtros definidos para os primeiros produtos

Produtos	Filtros
Leite	Tipo, Sabor e Marca
Água	Tipo, Sabor e Marca
Maçã	Tipo
Vinho	Tipo e Região

Após a definição dos filtros para estes produtos iniciais, foram então desenhadas as *prompts* do sistema de forma a auxiliar o utilizador a filtrar os resultados. Apesar de serem produtos diferentes, as questões eram iguais de acordo com os filtros como é possível consultar na Tabela 7.



Tabela 7. *Prompts* iniciais e gerais de acordo com o tipo de filtro

Filtros	Prompts
Tipo	<ul style="list-style-type: none"> <li>- Qual é o tipo de produto que deseja?</li> <li>- Qual é o tipo de produto que procura?</li> <li>- Tem preferência por que tipo de produto?</li> </ul>
Sabor	<ul style="list-style-type: none"> <li>- Prefere o produto com algum sabor?</li> <li>- Que sabor é que deseja?</li> <li>- Tem preferência por que sabor?</li> </ul>
Marca	<ul style="list-style-type: none"> <li>- Qual é a marca que procura?</li> <li>- Tem preferência por que marca?</li> <li>- Qual é a marca que prefere?</li> </ul>
Região	<ul style="list-style-type: none"> <li>- Tem preferência por que região?</li> <li>- Qual é a região que procura?</li> <li>- Qual é a região que deseja?</li> </ul>

De forma a ser possível obter uma maior variedade quanto aos filtros, foram assim definidos outros produtos, para além dos escolhidos anteriormente. A estes últimos foram realizadas algumas alterações quanto ao diálogo.

As *prompts* que envolvem o filtro quanto ao tipo de produto são as que mais se destacam nas alterações, pois dependendo do produto é possível questionar “Qual é o tipo de produto que prefere?” ou então para outros casos “Prefere o produto do tipo X ou Y?”. Esta última alternativa é mais indicada para produtos que tenham menos do que 4 tipos, sendo que a primeira é mais indicada para o oposto. Se fosse apresentada a segunda questão, para um produto com vários tipos, o utilizador não iria memorizar toda a informação devido à sua capacidade de memória e tornaria-se mais uma vez exaustivo.

Este mesmo processo de filtragem pode variar na ordem, consoante o produto em questão. Por exemplo, existem produtos nos quais os utilizadores dão maior importância a um determinado filtro do que a outro. Para tal, foram criados filtros mais específicos para produtos que necessitem, tal como é o caso do vinho. Neste, foi adicionado um para determinar o ano de preferência do utilizador quanto a este produto (Figura 34).

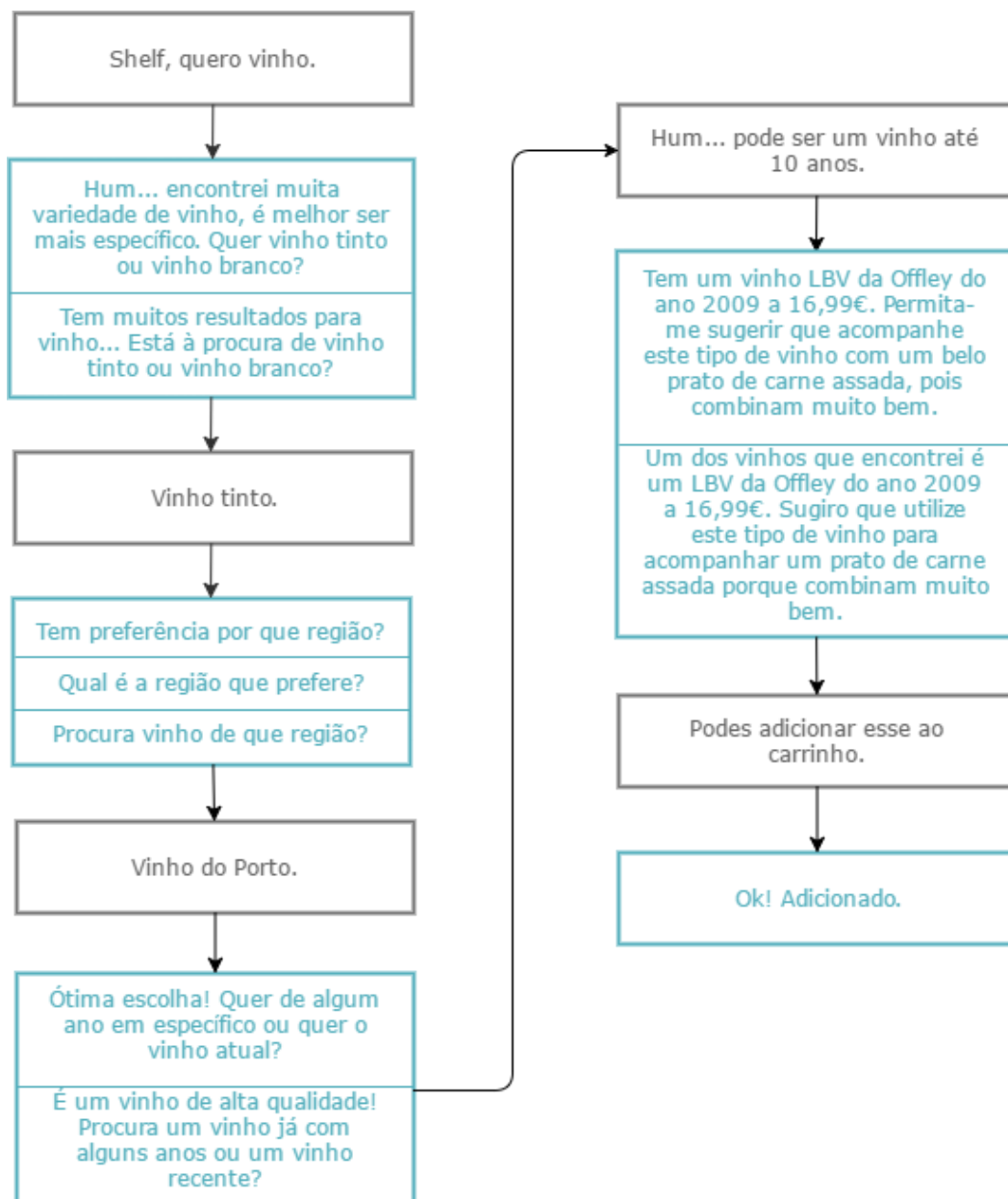


Figura 34. Exemplo de fluxo de diálogo para a pesquisa do produto vinho

Atualmente na aplicação Shelf AI, os produtos são ordenados com base na relevância, sendo que o primeiro é destacado dos restantes. Tendo esse aspeto em consideração, no momento em que o utilizador pesquisa um produto, foi definida a apresentação desse produto por voz. Essa apresentação inclui o seu nome, marca e preço e no caso de estar em promoção, é destacado esse facto (Figura 35).

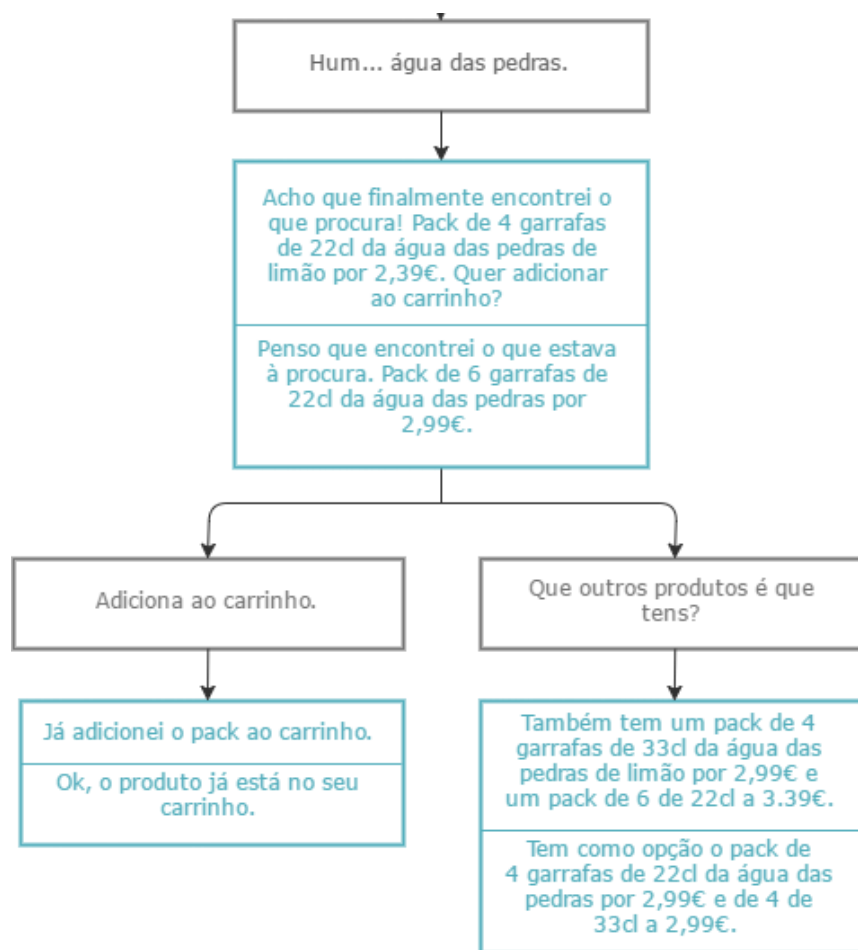


Figura 35. Exemplo de um momento de pesquisa filtrada e com resultados específicos

Quanto aos restantes produtos, se existirem mais resultados relacionados com o que o utilizador procura, não são apresentados por voz. Caso se trate de uma lista extensa, o utilizador não será capaz de memorizar todos os produtos e definir qual deseja. No entanto, se for expressamente pedido pelo utilizador para apresentar mais produtos semelhantes ao que ele procura, serão apresentados por voz os próximos dois produtos mais relevantes, tal como é possível consultar um exemplo no fim do diagrama da Figura 35.

A interface deve garantir uma adaptabilidade para utilizadores novatos e experientes, por isso foi decidido para a primeira utilização, providenciar ao utilizador uma espécie de tutorial com dicas de como interagir com a interface (Dybkaer & Bernsen, 2000), enquanto que na próxima utilização, o utilizador já não se depara com essas instruções. Deste modo, após a definição dos possíveis diálogos relativamente à pesquisa de produtos, foi notória a necessidade de criar uma ajuda inicial para os utilizadores entenderem como utilizar a interface.

Para tal, foram concebidas *prompts* de boas vindas que surgem com a primeira utilização do sistema. Posteriormente à mensagem inicial, para que o utilizador se sinta mais à vontade, foi

definida uma questão relativa ao nome do utilizador, simulando uma relação mais próxima e pessoal. De seguida, é questionado qual o produto que deseja comprar, para permitir que o utilizador compreenda que pode indicar de imediato o produto que deseja (Figura 36).

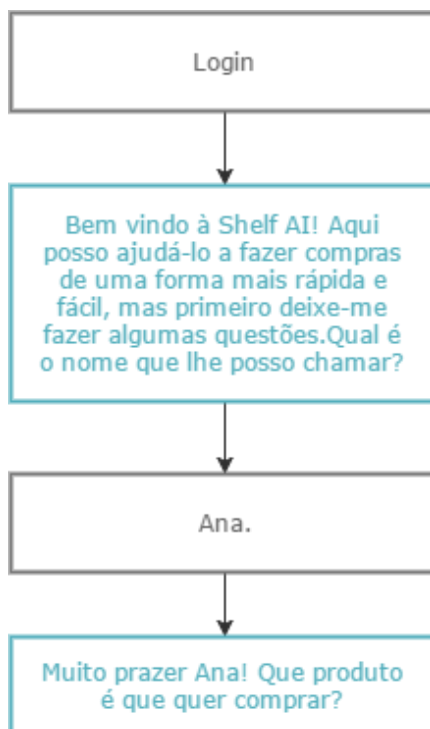


Figura 36. Fluxo de diálogo de boas vindas

Após três pesquisas realizadas com o auxílio dos filtros, é apresentada uma *prompt* para informar o utilizador de que pode tornar a sua pesquisa ainda mais fácil e rápida, de acordo com a quantidade de características que apresenta sobre o produto que procura (Figura 37).

" Como já percebeu, a forma mais eficaz para apresentar os produtos que procura é através de filtragem. Assim, pode dizer de uma só vez o máximo de características relacionadas com o produto, não sendo necessário fazer-lhe todas as questões."

Figura 37. Sugestão do sistema para uma pesquisa de produtos mais rápida

Após esta sugestão dada ao utilizador, este compreende que pode procurar o produto que deseja, especificando as características do mesmo, tal como é apresentado na Figura 38. Em relação à quantidade, se não especificar uma que deseja, será adicionada a quantidade por defeito, de uma unidade.

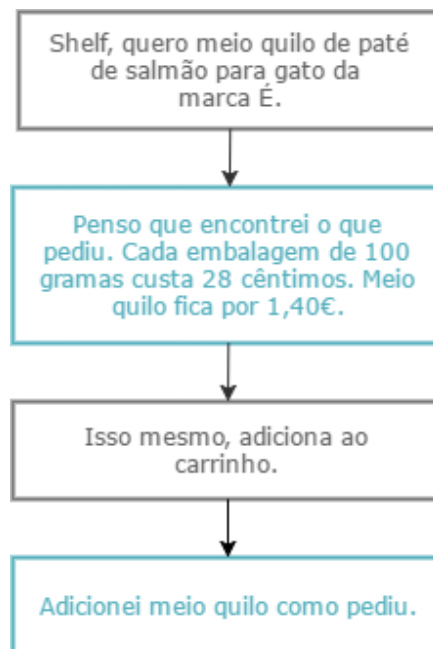


Figura 38. Exemplo de uma pesquisa completa em relação ao produto

Quando o utilizador, iniciante ou experiente, acede ao carrinho de compras, é reproduzida uma *prompt* informando quantos produtos tem e o total de custos da compra (Figura 39). É possível também a consulta de todos os produtos através do ecrã. Em relação ao número de produtos e o total de custos associados, estes são apresentados através de voz, pois a enumeração de todos os produtos existentes no carrinho de compras e no caso de uma lista extensa, seria mais uma vez, exaustivo para o próprio utilizador. Desta forma, apenas são apresentados todos os produtos por voz, caso seja explicitamente requisitado pelo utilizador.

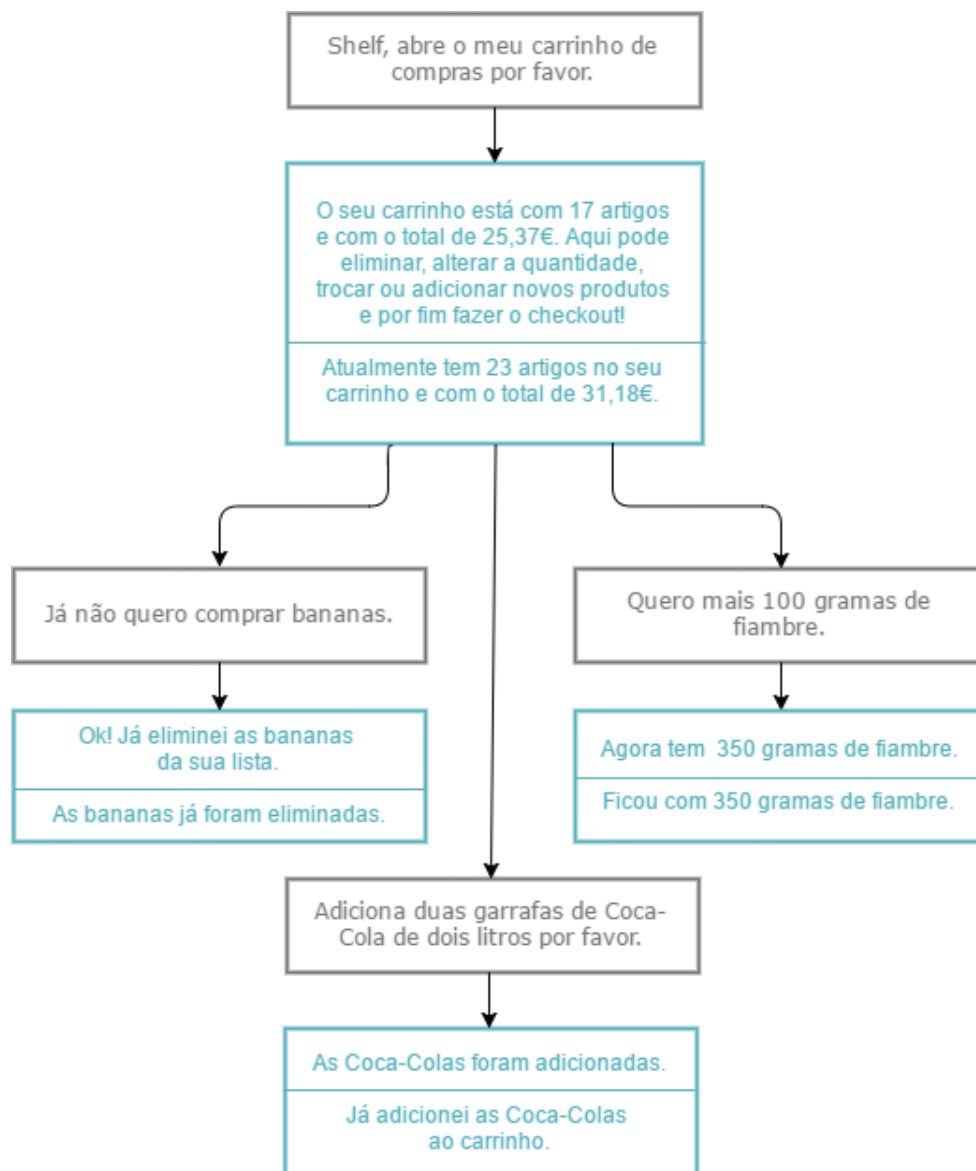


Figura 39. Interação entre o utilizador e o sistema no carrinho de compras

Como forma de auxiliar o utilizador na marcação da entrega, foi idealizado um diálogo para o *checkout*. Nesta etapa são pedidos os dados comuns como o nome, morada de entrega, dia e hora de entrega e escolha do método de pagamento. No caso da marcação do dia e hora, o utilizador pode deparar-se com uma *prompt* informando que para esse dia e hora escolhidos não é possível a entrega e de seguida é apresentada uma sugestão. Caso o utilizador não aceite, será apresentada novamente outra sugestão e se mais uma vez não aceitar, o utilizador fica livre de tentar marcar através da aplicação. É também sugerido aceder ao calendário da Shelf AI, onde pode consultar os horários e dias disponíveis. Após ser pedido o método de pagamento que o utilizador prefere, a interface sugere que os dados sejam introduzidos manualmente por questões de segurança. Por fim, a interface repete toda a informação inserida em relação à encomenda, para que o utilizador possa confirmar se está tudo correto e assim finalizar o processo de encomenda.

Em relação ao estado de encomenda, e para que o acesso a esta informação seja facilitado, existe a possibilidade de o utilizador questionar diretamente à Shelf AI ou então obter mais informações através do email. Todo o processo de *checkout* abordado pode ser perceptível através do diagrama presente na Figura 40.

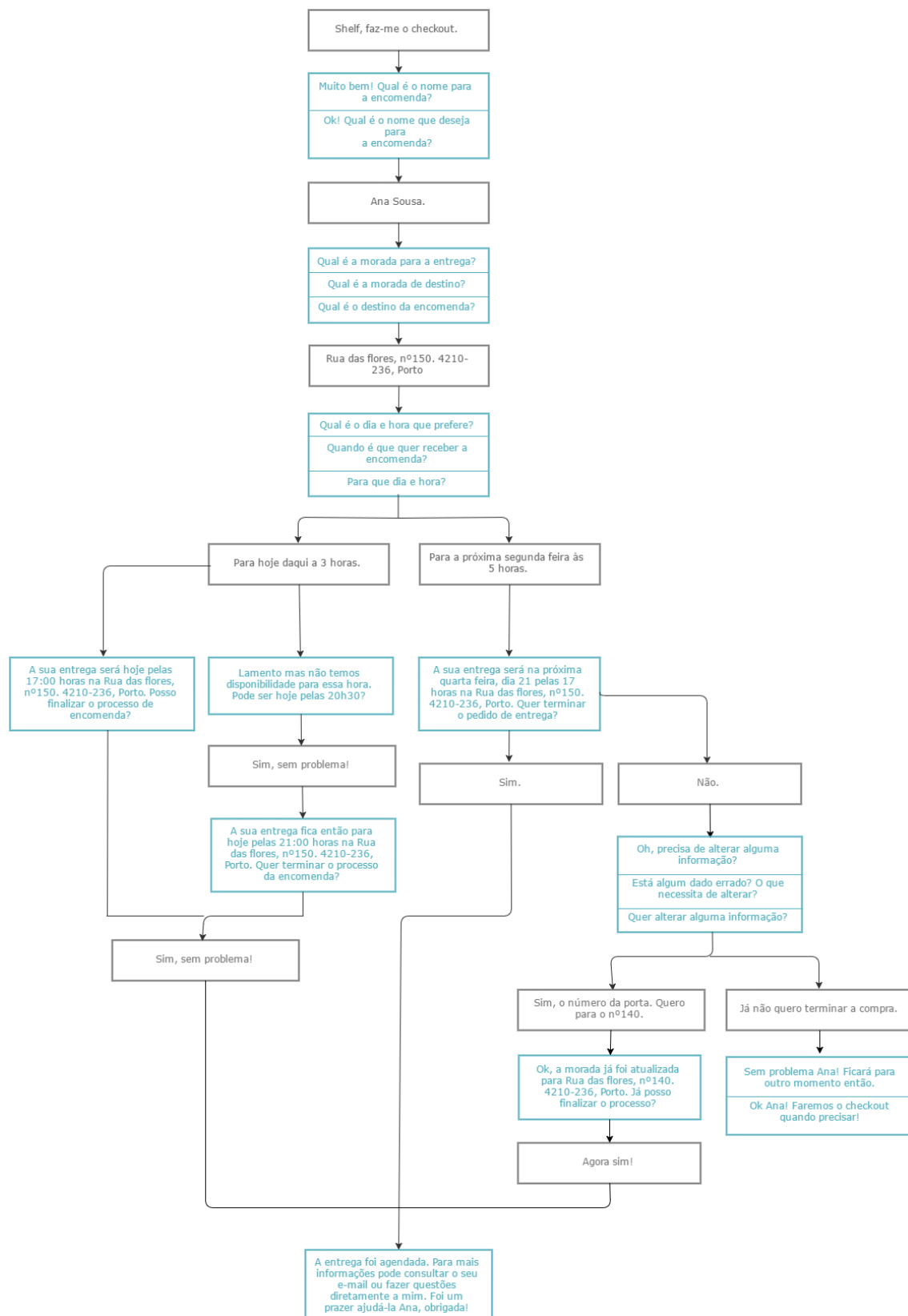


Figura 40. Processo de *checkout* com diferentes tipos de interação e escolhas



Quando o utilizador pretende tirar dúvidas associadas a produtos, as respostas são apresentadas por voz, mas em simultâneo o utilizador pode encontrar os resultados no ecrã. Por exemplo, se o utilizador perguntar “Qual é o café da marca X com mais cafeína?”, a resposta da do sistema será “O café com mais cafeína é o Y, mas se não for o que procura, poderá ver os resultados ordenados por ordem decrescente de acordo com o nível de cafeína”. Pode-se conferir através da Figura 41 que é possível apresentar outras questões à interface para obter respostas no mesmo momento.

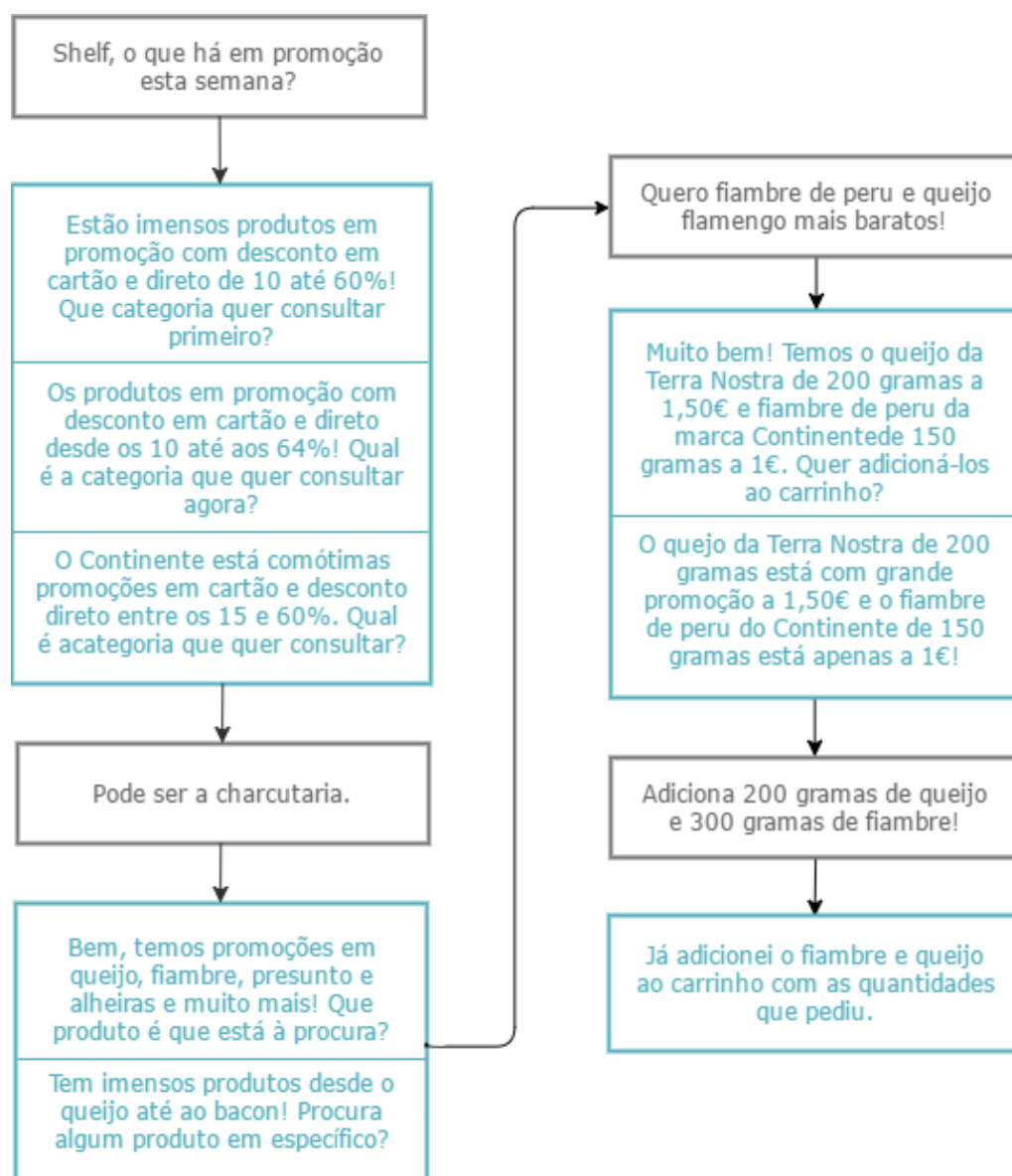


Figura 41. Dúvida por parte do utilizador com respostas associadas

Sempre que decorre algum problema de compreensão, o sistema deve agir consoante a situação. Por exemplo, se o utilizador não responder a uma questão, o sistema deve informar que

está a aguardar a entrada. Se o utilizador rejeitar o reconhecimento do sistema, tem a oportunidade de corrigir.

Usufruindo da componente de inteligência artificial e de modo a otimizar o processo de personalização de acordo com o utilizador, foi criada uma regra para que o sistema armazene a informação e utilize-a nas futuras pesquisas. Essa regra está relacionada com as palavras “sempre” e “nunca”.

Se a entrada do utilizador for algo como “Quero sempre bolachas sem açúcar”, o sistema vai identificar o “sempre” e a categoria do produto pedido. Assim que o utilizador procurar por “bolachas”, o sistema sabe que é provável que este está especificamente à procura das bolachas sem açúcar.

Já no caso da entrada do utilizador referir uma característica específica a um produto, não contando com os atributos gerais como marca, sabor, tipo, entre outros, o sistema também deve armazenar essa informação. Por exemplo, se o utilizador referir que quer um creme de corpo para pele seca, o sistema memorizará essa preferência pelo produto, mas também a característica “pele seca” em separado. Assim, procurando mais tarde por exemplo, por “creme hidratante para rosto”, os primeiros resultados da lista serão cremes hidratantes para rosto do tipo de pele seca.

De modo a que essas características importantes sejam memorizadas e de forma a personalizar os resultados de acordo com o utilizador, nas primeiras interações foram desenhadas *prompts* para esses casos. Essas *prompts* têm o objetivo de mostrar ao utilizador que o sistema armazena as suas preferências e necessidades, como é possível conferir na Figura 42.

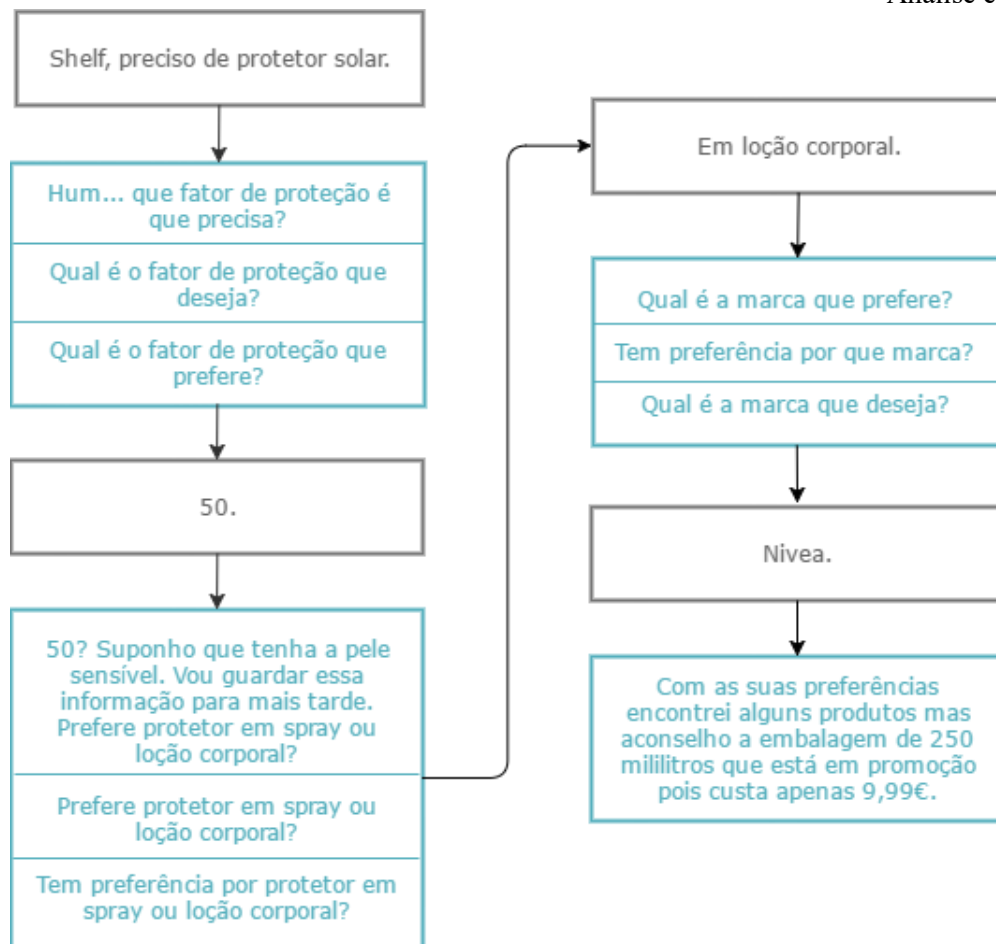


Figura 42. Fluxo de pesquisa que permite detetar informação importante a armazenar

Tal como é possível conferir na figura, após o utilizador escolher o fator de proteção 50, o sistema mostra ao utilizador, através da *prompt*, que está a armazenar a informação que considera importante para personalizar os resultados da pesquisa.

### 3.7. Síntese do Capítulo

Neste capítulo foi apresentado o estudo realizado antes de ser iniciada a definição e concepção do *design*, assim como a definição da aplicação em relação ao público-alvo, identificação dos fatores diferenciadores, as funcionalidades e fluxos de interação.

Após a análise de um estudo estatístico do INE e tendo em consideração as respostas obtidas no questionário inicial, definiu-se que o público-alvo da Shelf AI terá idades compreendidas entre os 18 e 55 anos. Esse mesmo questionário ajudou a traçar mais características do perfil do público-alvo, definindo que os utilizadores podem ser pouco ou muito experientes com *smartphones*, pois podem interagir por voz, e podem ou não ter o hábito de realizar compras *online*.

O questionário também permitiu compreender que poucos participantes já utilizaram uma interface conversacional de voz, enquanto que grande parte já utilizou *chatbots* e outros tantos nunca utilizaram nenhuma interface conversacional. A maioria dos participantes prefere que tanto a interface, como eles enquanto utilizadores, possam interagir de forma híbrida. No entanto, também foi notório um grande favoritismo pela interação do utilizador com a interface através do toque/texto.

Depois de uma análise a interfaces associadas ao *e-commerce*, reconheceu-se que o principal ponto de diferenciação da Shelf AI, em relação a estas, é o facto de ser uma interface conversacional híbrida, enquanto que as outras resumem-se a interfaces gráficas com pouca inteligência artificial ou são interfaces unicamente por voz. Posteriormente foram definidas as funcionalidades que referem o ponto de vista do utilizador ao interagir com o sistema e também referentes ao ponto de vista do sistema ao comunicar com o utilizador.

Após a definição destes pontos, foi possível iniciar a fase do *design* de usabilidade conversacional. Inicialmente foi analisada e estudada a Shelf AI original, de forma a compreender quais as alterações podiam ser feitas, essencialmente em relação a quando e a como adicionar o diálogo na aplicação. A Shelf AI foi definida como sendo um sistema de iniciativa mista, e concluiu-se que o sistema deve apresentar um *feedback* ao utilizador do que está a ocorrer durante a interação. Um exemplo deste *feedback* foi a inserção de *pop ups* associados ao reconhecimento de voz, de modo a informar o utilizador de que o sistema está realmente a reconhecer o diálogo e que está a processar nesse momento.

O *checkout* também foi alterado, criando um ecrã onde é necessário preencher todos os campos com a informação necessária, em alternativa a aceder ao *site* do retalhista para realizar o *checkout*, visto que seria um processo mais demorado e obrigaria o utilizador a trocar de plataforma.

O principal foco desta fase e o ponto mais importante da interface está relacionada com o *design* do fluxo do diálogo. Para desenhar o fluxo de diálogo foi necessário ter em conta as funcionalidades definidas para desenhar os diálogos de acordo com cada uma. Também foi essencial refletir sobre a forma como os utilizadores procurariam pelos produtos, como o sistema pode interagir de uma forma mais natural e quais os momentos em que não é favorável a interação por voz por parte do utilizador, assim como pelo sistema.

Durante a interação podem existir determinados momentos em que não é vantajoso a interface comunicar por voz, tal como acontece caso seja feita uma pesquisa em que surjam mais do que dois resultados ou quando é feita a consulta do carrinho de compras que inclua mais do que três produtos, tal como referido anteriormente. Em consequência, decidiu-se que a interface só apresentará os produtos da pesquisa em lista por voz, se o utilizador assim o pedir.

Para a pesquisa de produtos, foram desenhadas várias *prompts* para haver filtragem, de forma a que a experiência não se torne repetitiva e exaustiva para o utilizador, após várias pesquisas consecutivas. Essas mesmas *prompts* de filtragem podem variar de tipo consoante o produto em questão, assim como a ordem de filtros. Para uns produtos existem filtros que para outros não fazem sentido momentos favoráveis ou desfavoráveis para a interação por voz, tentando resolver da melhor forma as possibilidades de interação para esses momentos. Existem vários momentos desfavoráveis que foram trabalhados neste projeto, como é o caso do momento em que existe uma grande lista de produtos e também quando é necessário apresentar informação confidencial e importante.

## 4. Definição dos Protótipos e Testes

Neste capítulo será apresentado todo o processo de realização dos protótipos, desde os desafios e limites das tecnologias para realizar protótipos de interfaces conversacionais híbridas, assim como quais as soluções encontradas para executá-los. Será também abordada a fase de testes onde são definidos os objetivos em realizá-los, uma análise do perfil dos participantes e as tarefas a serem executadas nos testes. Por fim, é definido o questionário a ser respondido pelos participantes após os testes, uma análise dos resultados recolhidos durante estes, assim como dos dados obtidos com os questionários.

### 4.1. Desafios nas Tecnologias

No início da fase de realização dos protótipos, foram detetados vários problemas relacionados com as limitações das tecnologias que podem ser utilizadas, sendo necessário contorná-los com recurso a outras soluções, através da descoberta de ferramentas alternativas para a realização de protótipos e de um método ideal para os testes de usabilidade.

Após uma vasta pesquisa de ferramentas de prototipagem para criar protótipos de interfaces conversacionais híbridas, concluiu-se que não existe nenhuma ferramenta ideal para este tipo de interfaces. Para satisfazer a necessidade de conversação foram encontradas várias ferramentas *online* (Tabela 8).

Tabela 8. Ferramentas para simular e/ou implementar conversações

Ferramenta	Descrição
SaySpring	A SaySpring dá a possibilidade de realizar testes a interfaces de voz sem recorrer à programação. Permite criar diretamente fluxos de diálogo com o intuito de testá-los através da Amazon Alexa ou Google Assistant. A ligação ao Google Assistant é efetuada através de um comando e com a Alexa através de uma <i>skill</i> com o nome Sayspring, disponível na Alexa Skills.
Wit.ai	Comprada pelo Facebook, a Wit.ai é composta pela habilidade de reconhecimento de voz e <i>machine learning</i> . Permite ao utilizador interagir com <i>bots</i> através de texto e voz, assim como ter acesso a uma interface de voz numa aplicação através de um <i>smartphone</i> ou de <i>wearable devices</i> .
Api.ai	A Api.ai, comprada pela empresa Google é muito semelhante à Wit.ai. Permite a construção de interações por linguagem natural para aplicações, serviços, <i>chatbots</i> ou interações por voz para dispositivos como o Google Home ou Amazon Echo, conectando-se assim à Alexa. Deste modo, o utilizador tem a possibilidade de interagir em diferentes dispositivos através de texto e voz.
ChatScript	O ChatScript é uma ferramenta que para além de ser um motor de processamento de linguagem natural, serve também como gestor de diálogos. Inicialmente esta ferramenta tinha o objetivo de criar <i>chatbots</i> , mas entretanto também passou a ser utilizada noutras formas de processamento de linguagem natural.

Estas ferramentas *online*, permitem a criação de fluxos de diálogo onde é possível testá-los sem necessidade de recorrer à implementação. Eventualmente se o objetivo fosse implementar, as ferramentas utilizadas seriam a Wit.ai ou Api.ai, pois a Sayspring não passa de uma ferramenta de testes para interfaces conversacionais. Esta ferramenta foca-se exclusivamente na criação de fluxos de diálogos que concebem o percurso de interação do utilizador e que têm a função de organizar o grupo de pedidos do utilizador, bem como as respostas da interface.

O ChatScript é uma ferramenta que possibilita implementar diálogos ideais para *chatbots*. Os objetivos da presente dissertação afastam-se da implementação, contudo, é de salientar que se poderia recorrer a esta ferramenta para implementar o diálogo na Shelf AI. Um dos pontos interessantes é que o ChatScript é ideal para momentos de *helpdesk*, isto é, de apoio ao cliente,

pois mesmo depois de implementado, as suas ações podem ser um pouco auxiliadas por agentes humanos.

Quanto à Wit.ai e Api.ai, estas fornecem também a possibilidade de produzir diálogos, no entanto, não se limitam à criação de um fluxo, uma vez que é necessário serem definidas diversas componentes, com o intuito de auxiliar o processamento de linguagem natural e a inteligência artificial.

Seria ideal inserir na Api.Ai os fluxos de diálogo em desenvolvimento, contudo seria necessário recorrer à programação para conectar a componente da voz, na aplicação da Shelf.Ai, assim como para aceder à base de dados dos produtos existentes para obter os resultados de pesquisa. Este passaria de um protótipo para uma implementação aproximada da final, não sendo possível na presente dissertação.

Para criar e desenhar uma interface conversacional de uma forma mais correta, é importante compreender as tecnologias que são utilizadas para construir as interações com base em linguagem natural. Posteriormente, surgiu a necessidade de perceber quais as limitações das tecnologias, que por sua vez podem ser utilizadas para auxiliar na criação e alteração do *design* dos fluxos de diálogo.

Para tal, foi iniciado um estudo da Api.ai através da leitura da documentação para saber como inserir os fluxos de diálogo na ferramenta. Para esses fluxos serem possíveis e funcionais, eram necessárias as definições dos vários componentes, sendo que os principais utilizados foram as entidades e os *intents*.

Nas entidades foram definidas as listas de produtos, características e marcas. Essa lista tem como finalidade determinar que o sistema analisa e deteta se a entrada do utilizador contém algum desses elementos, para assim reconhecer e agir.

Nos *intents* foram definidos diversos elementos: a entrada do utilizador, ações, contextos, e repostas do sistema. Na entrada do utilizador foram demarcados vários exemplos das frases e as entidades foram associadas a determinadas palavras dessas entradas.

As ações estão relacionadas com o que o sistema vai fazer em função da entrada do utilizador e podem ser definidas após e de acordo com as entidades associadas a cada palavra. Essas ações podem ser marcadas como obrigatórias, significando que para chegar ao fim do *intent* é necessário percorrer essas ações.

Os contextos têm como objetivo interligar os *intents*, no qual um *intent* pode receber vários contextos de outros, assim como iniciar e exportar contextos para outros. A criação e utilização de contextos é vantajosa para diversos momentos de interação com a Shelf AI. Por exemplo, no caso da pesquisa filtrada de produtos, é necessário o sistema armazenar qual é o produto e as características já escolhidas anteriormente, de forma a filtrar e apresentar resultados correspondentes a essas características.



Por fim, as respostas do sistema são desenhadas para surgirem de acordo com o *intent* em questão e assim responder à entrada do utilizador. Nessas mesmas respostas podem estar presentes as entidades através dos nomes dados a essas.

Com isto, foi possível compreender como esta tecnologia funciona e quais são os seus limites e cuidados a ter, para não existirem quebras de fluxo e de interação. Posteriormente, foi possível continuar o processo de *design* do fluxo de diálogo, de forma a que o diálogo seja estruturado cuidadosamente, tendo em conta como as tecnologias de implementação funcionam.

Concluindo, as ferramentas como a Wit.Ai, Api.Ai e ChatScript podem ser utilizadas para implementar na Shelf AI a componente de diálogo natural. Com a utilização destas, é necessário recorrer também a um motor de síntese de voz, Text to Speech (TTS) para a interface interagir por voz, assim como um Speech to Text (STT) para o sistema converter a fala do utilizador em texto. Posto isto, o TTS está por exemplo ligado à Api.Ai e a partir daí é extraído o objetivo que inclui o *intent* e por exemplo, a lista de resultados do produto.

## 4.2. Tecnologias e Métodos Utilizados

Após uma análise do problema, do projeto, do estado de arte relacionado com o *design* de interfaces e das ferramentas de prototipagem para realizar interfaces conversacionais, foi possível obter várias conclusões.

Para a realização dos testes de usabilidade, foi necessária a criação de um protótipo interativo e estático. Não existindo nenhuma ferramenta ideal para a criação de protótipos para interfaces híbridas, foi necessário encontrar outra solução. O protótipo foi desenvolvido através da plataforma Marvel, onde é possível realizar protótipos de interfaces gráficas que funcionam através de cliques. A escolha desta ferramenta deve-se ao facto de já ter sido utilizada em projetos realizados no Mestrado em Multimédia, que levaram a bons resultados. Além da experiência prévia, a ferramenta atualmente apresenta atualizações e novidades nas funcionalidades, podendo assim acelerar e melhorar o processo de realização do protótipo.

Uma vez que o protótipo foi desenvolvido nesta plataforma, era composto apenas pela interação por toque, surgindo assim a necessidade de encontrar uma solução para testar a interação por voz do utilizador com a interface, assim como da interface com o utilizador. Após uma análise às diversas possibilidades de testes e tendo em conta que não existem ferramentas de protótipos para interfaces como esta, capaz de simular o reconhecimento de voz e gerar respostas automáticas sem recorrer à programação, concluiu-se que o método mais indicado para realizar os testes seria o método de WoZ. Este método consiste na realização de testes em que o seu funcionamento é simplesmente simulado e controlado pelo responsável do teste. O seu objetivo é providenciar uma experiência de teste que pareça real, ou seja, que o sistema está

a funcionar de forma independente. O WoZ foi essencialmente escolhido para simular a interação por voz com a interface.

Com a utilização deste, o *designer* pode criar os cenários de uso e preparar a lista de tarefas que o utilizador deve realizar, de forma a que possa preparar previamente todas as *prompts* necessárias para interagir com o utilizador durante as tarefas. Além dessas *prompts*, também é oportuno criar outras para diferentes contextos de uso, como por exemplo uma de boas vindas, de instruções, de erro, de falta de reconhecimento caso exista um desvio por parte do utilizador, da entrada necessária para realizar a tarefa. Estas serão gravadas previamente e então organizadas consoante as tarefas, de modo a que o *designer* possa selecionar rapidamente a *prompt* que corresponde à entrada do utilizador.

Para simular o reconhecimento das entradas de voz do utilizador, foram criados e inseridos *pop ups* no protótipo, com um ícone de som, juntamente com uma frase semelhante à que o utilizador podia dizer consoante cada tarefa. Tal como referido no subcapítulo anterior, os *pop ups* surgem de acordo com a resposta do utilizador e para isso, foi necessária a criação de *botões* invisíveis para o *designer* clicar e assim surgir a mensagem, sem perturbar a experiência. De acordo com a escolha do utilizador, foram interligadas as mensagens de reconhecimento de voz aos ecrãs dos resultados, consoante a escolha anterior.

Para as *prompts* da interface, foi necessário recorrer a um TTS sendo utilizado um *online* da Oddcast – Character Driven Communication. Devido a certas limitações de leitura derivantes do TTS escolhido, foi necessário reformular no momento de gravação de áudio as *prompts* já desenhadas. Algumas alterações nas pontuações e palavras foram fundamentais, de modo a melhorar a leitura das frases e torna-la o mais natural possível. Após a reformulação e gravação das *prompts*, utilizando o *software* Audicity, estas foram devidamente identificadas e estruturadas para facilitar ao *designer* a reprodução, de acordo com as ações do utilizador. Essa ação, no momento certo, contribuiu para a simulação da inteligência artificial e da capacidade de conversação que existe nas interfaces.

Por fim, foi possível o controlo remoto recorrendo à ferramenta TeamViewer no computador, que se conectava ao *smartphone* através da aplicação TeamViewer QuickSupport. Isto permitiu a realização dos testes com o método de WoZ. Através desta ferramenta de controlo remoto, o protótipo realizado na Marvel era controlado em alguns momentos da interação com *botões* invisíveis. Quanto à reprodução das *prompts* do sistema, estas ocorriam através da *stream* realizada, recorrendo ao uso do *software* SoundWire Server, possibilitando a reprodução a partir do computador para o *smartphone*.

### 4.3. Testes de usabilidade

Os testes de usabilidade são um método eficaz para melhorar e corrigir os possíveis problemas existentes numa aplicação, através da observação da interação do utilizador com esta. Segundo Cuello e Vittone (2013), estes devem ser realizados nas fases iniciais do projeto antes de passar para a implementação da versão final do mesmo.

Na fase de testes foram inicialmente definidos os objetivos dos testes de usabilidade, para delinear o guião do teste, que inclui as tarefas e então passar para a escolha dos participantes. Durante os testes, os utilizadores foram observados de forma a analisar o seu comportamento e interação. Essa análise foi auxiliada com o recurso às filmagens realizadas nos momentos de interação.

Com isto tornou-se possível compreender melhor os problemas existentes e medir a usabilidade, tendo em consideração os cinco atributos que Nielsen (2012) refere. Esses atributos são baseados na facilidade de aprendizagem com que o utilizador se depara ao interagir com o sistema, mesmo sendo um novato, na eficácia da interface dependendo da forma de interação do utilizador, na fácil memorização da interface que depende da eficácia com que o utilizador interage com esta e a fácil recuperação de erros. O quinto atributo é o que se encontra mais relacionado com os objetivos da usabilidade, baseia-se no grau de satisfação do utilizador, onde este pode ser afetado pela frustração, dependendo dos resultados anteriores e do *design* geral da interface.

A informação a ser obtida nos testes, teve como base três componentes definidos por Nielsen (2012), onde é necessário reconhecer os representantes do público-alvo, apresentar tarefas representativas da interface, para serem realizadas pelos utilizadores, assim como observar a interação do utilizador com base nos passos que davam e resultados obtidos das tarefas.

Recorrendo ao método de teste de WoZ com cenários de tarefas representativas, foi possível obter dados importantes sobre a interface testada em protótipo. As *prompts* do sistema devem ser analisadas com rigor, de forma a identificar sinais de que o utilizador está com problemas em relação às capacidades do sistema.

Por fim, foram também realizados questionários após o teste, com a principal finalidade de tomar conhecimento das opiniões dos utilizadores em relação à interface testada, assim como qual o grau de satisfação dos mesmos.

#### 4.3.1. Objetivos dos Testes

Estes testes tinham objetivos que passavam por avaliar a eficácia da interface híbrida para compras, a adequação das *prompts* do sistema, a utilidade da conversação por voz em relação à

sua usabilidade, problemas de compreensão por parte do utilizador e a adaptação do utilizador à interação por voz, assim como a sua preferência de interação.

Os principais pontos a serem avaliados no teste, através da análise da interação do utilizador com a interface, consistiam em três pontos fundamentais :

1. Momentos de dúvida ao interagir – hesitação e tempo de reação do utilizador;
2. Qual o tipo de pesquisa preferida – voz ou texto/manual;
3. Diálogo natural ou mecânico por parte do utilizador;

Em adição, foi também realizado um teste com a aplicação original Shelf AI, com as mesmas tarefas, de forma a que o utilizador pudesse dar a sua opinião em relação a qual interface era a mais útil para realizar compras e qual seria a sua escolha entre essas.

#### **4.3.2. Perfil dos participantes**

De acordo com Nielsen (2012), um dos componentes a ter em conta na fase dos testes passa por reconhecer os representantes do público-alvo. As informações que resultaram da amostra dos 11 participantes dos testes de usabilidade, mostraram que os utilizadores têm idades compreendidas entre os 20 e 34 anos, em que todos possuem experiência e contacto diário com *smartphones* e aplicações. A maioria dos participantes são estudantes, dos quais 2 vivem temporariamente fora da sua zona de residência natural e 3 são trabalhadores. Por fim, 8 dos participantes já realizaram compras *online* e apenas 3 já tiveram contacto com interfaces conversacionais.

#### **4.3.3. Tarefas**

Para os testes de usabilidade foram definidas várias tarefas representativas, que vão de encontro aos componentes definidos por Nielsen (2012), para explorar diferentes momentos de interação e funcionalidades, assim como delimitar a interação necessária para a realização dos protótipos. Foi realizado um guião de teste de usabilidade onde estão incluídas as tarefas e que é possível consultar no Anexo A.

No início do guião é apresentado o âmbito do projeto, através de uma síntese que explica em que consiste uma interface conversacional híbrida. De seguida, é apresentado o comando por voz, que é um ponto fundamental no teste, para o utilizador entender que deve recorrer ao mesmo para ativar o modo de interação por voz quando não está a decorrer nenhum momento de diálogo. Por fim, foi evidenciado o facto de o teste ser baseado num protótipo, com uma breve explicação sobre este, para que o utilizador possa ter consciência das potenciais limitações com que se pode deparar.

O teste realizado tinha como objetivo simular a primeira interação do utilizador na interface, mostrando que há uma evolução ao longo da realização deste, por conseguinte foram definidas sete tarefas para esse fim (Tabela 9).

Tabela 9. Tarefas do teste de usabilidade

Tarefas
1. Faça login na aplicação através do facebook como se o seu nome fosse Diana ou Ricardo. Quando surgir oportunidade pesquise por “batatas fritas”.
2. Pesquise por “protetor solar” e após concluir a filtragem, adicione ao carrinho o produto destacado.
3. Pesquise por “vinho” e no fim da pesquisa, adicione ao carrinho o primeiro produto.
4. Pesquise por “iogurte sem lactose de morango” e após concluir a pesquisa, adicione ao carrinho o produto de topo.
5. Pesquise por “água das pedras de limão” e adicione ao carrinho o primeiro produto.
6. Descubra qual é a compota com menos calorias.
7. Realize o checkout e defina que a morada de entrega na Rua das flores, número 1. Marque a entrega para a próxima terça feira pelas 15h00. Caso não seja possível tente manter a mesma hora para outro dia.

Na primeira tarefa, o utilizador tinha de efetuar o *login* e de seguida escolher o nome pelo qual queria ser tratado pela interface, de forma a familiarizar-se com a capacidade de existir um diálogo natural.

As três tarefas seguintes consistiam na pesquisa de produtos genaralizados, de modo a que fosse necessário o sistema fazer questões ao utilizador para filtrar os resultados. Essas questões consistiam por exemplo, no tipo de produto que o utilizador preferia, se procurava algum sabor em específico, entre outros. Deste modo, é possível filtrar os resultados de pesquisa até diminuir ao máximo o número de resultados presentes, o que facilita a descoberta e a escolha do produto desejado.

Após as três primeiras pesquisas, foi apresentada uma sugestão ao utilizador, informando-o de que tinha a possibilidade de pesquisar um produto com mais características, para chegar de uma forma mais rápida e fácil ao resultado desejado. Assim, as duas tarefas seguintes consistiam na pesquisa de produtos com mais detalhe, como por exemplo “iogurte sem lactose de morango”, faltando apenas o filtro da marca. No caso de incluir na pesquisa todos os atributos de um produto, tal como ocorre com “água das pedras de limão”, não é necessário recorrer a nenhum filtro, passando de imediato para a lista final de resultados.

A penúltima tarefa tinha como finalidade o utilizador retirar uma dúvida sobre determinado produto, sendo neste caso necessário questionar qual a compota disponível com menos calorias. Esta tarefa permitiu apresentar ao utilizador a funcionalidade que consiste em retirar dúvidas de uma forma rápida sem recorrer à interação por toque.

Por fim, o utilizador tinha como tarefa realizar o *checkout*, onde na marcação da entrega se depara com um diálogo para definir uma data disponível, visto que a introduzida inicialmente se encontrava ocupada. Este pormenor possibilitou apresentar ao utilizador a utilidade que uma interface híbrida pode ter para resolver problemas momentâneos.

#### **4.3.4. Resultados**

Com os testes de usabilidade realizados, foi possível obter certas informações e conclusões em relação ao *design* da interface conversacional híbrida e também ter como base os três pontos fundamentais a serem avaliados durante os testes, como referido no subcapítulo 4.3.1.

Na primeira parte da tarefa, quando foi pedido o nome pelo qual o utilizador queria ser tratado, alguns participantes mostraram hesitação em como podiam responder, no entanto, todos acabaram por responder através de voz. Supõe-se que esta hesitação esteja relacionada com o facto de ser o primeiro momento de interação com a interface, ficando o utilizador um pouco na dúvida de como interagir.

Apenas dois utilizadores interagiram de forma muito mecânica por voz, sendo notável que as entradas limitavam-se ao nome do produto, assim como em outras ações não utilizaram frases naturais. Os outros utilizadores apresentaram um diálogo natural, com a formulação de frases em alguns momentos da interação.

Na tarefa 6, onde foi pedido ao utilizador para descobrir qual a compota com menos calorias, era de esperar que a maioria dos utilizadores regissem com uma entrada natural, formulando assim uma questão para a interface. No entanto, apenas 3 utilizadores fizeram uma questão, enquanto que 7 limitaram-se a dizer o que era pedido na tarefa – “compota com menos calorias” – e 1 utilizador pesquisou apenas por “compota”.

Todos os utilizadores adicionaram os produtos ao carrinho de compras manualmente, à exceção de 3 utilizadores que só nas três primeiras tarefas adicionaram manualmente e depois

passaram a adicionar por voz. Dois desses utilizadores questionaram se podiam adicionar por voz, enquanto que o outro arriscou em explorar essa opção sem saber que existia.

Na última tarefa, onde era pedido para realizar o *checkout*, 5 utilizadores entraram na zona de *checkout* através do carrinho de compras, enquanto que os restantes 6 entraram por voz. Os utilizadores na sua maioria, optaram por estruturar uma frase. Alguns exemplos foram: "Quero fazer o *checkout*.", "Quero pagar." e "Quero finalizar a compra". Quanto ao preenchimento dos dados necessários para o *checkout*, 9 dos utilizadores introduziram os dados por voz, enquanto os 2 restantes por texto. Para terminar o *checkout* 2 utilizadores escolheram terminar o processo com recurso ao botão "terminar", enquanto que os restantes terminaram por voz, onde 7 responderam com "sim" à *prompt* que questiona se pode finalizar o processo de encomenda, e os restantes 2 utilizaram as palavras "terminar" e "finalizar".

#### 4.3.5. Questionário Pós-Teste

Após a realização dos testes de usabilidade, foram efetuados questionários aos utilizadores, com o objetivo de avaliar eficazmente o seu nível de satisfação, suplementando assim a informação recolhida nos testes. Tal como defendem Dybkjær & Bernsen (2001), o questionário é um bom método para complementar as informações recolhidas nos testes e analisar o grau de satisfação dos utilizadores.

O questionário, disponível para consulta no Anexo B era composto por duas secções, uma com questões quantitativas e outra com questões de resposta aberta.

A primeira secção tinha cinco perguntas quantitativas e eram compostas por uma escala de 1 a 5, em que 1 significava “não concordo” e o 5 “concordo plenamente”. Já a segunda era composta por quatro perguntas qualitativas que permitiam uma resposta aberta. É possível verificar essas mesmas questões na Tabela 10.

Tabela 10. Questões quantitativas da primeira secção

Questões
1. Considera que a linguagem do sistema, em relação à estrutura das frases, foi natural/humana?
2. Considera útil existir uma espécie de tutorial integrado na primeira interação do utilizador com a interface? Tal como ocorreu no teste no momento da filtragem de produtos?
3. Gostou das funcionalidades existentes (sugestões de produtos, ajuda no checkout, personalização de pesquisa, respostas a eventuais dúvidas)?
4. Considera esta interface híbrida vantajosa em relação a uma interface unicamente gráfica?
5. Considera a interação por voz oportuna para certos contextos e situações?

Uma das questões tinha como finalidade obter a opinião dos utilizadores quanto ao diálogo natural da interface, questionando assim se consideravam a linguagem do sistema em relação à estrutura das frases de carácter natural e humano. Todos concordaram que a linguagem é natural, sendo que desses, 63,6% (7 participantes) concordam plenamente (Figura 43).



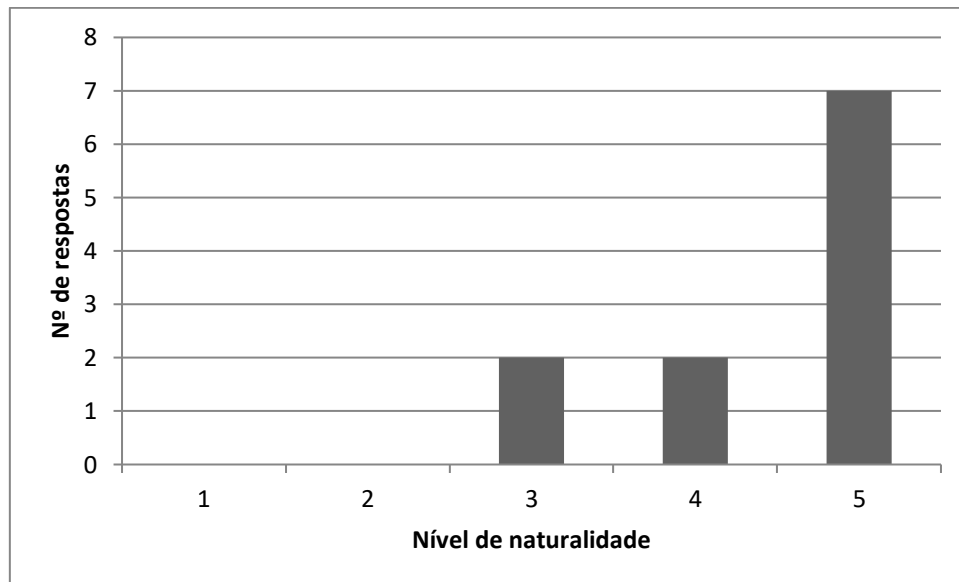


Figura 43. Nível de naturalidade da fala do sistema

Tendo em conta que no teste, o objetivo era a interação sofrer uma evolução de aprendizagem, adequando-se ao número de interações, pretendeu-se obter opinião dos utilizadores quanto à utilidade que vêm na integração de uma espécie de tutorial na primeira utilização do sistema. O tutorial é visto como algo muito positivo pois 90,1% dos utilizadores (10) concordaram plenamente que esses tutoriais podem ser úteis (Figura 44).

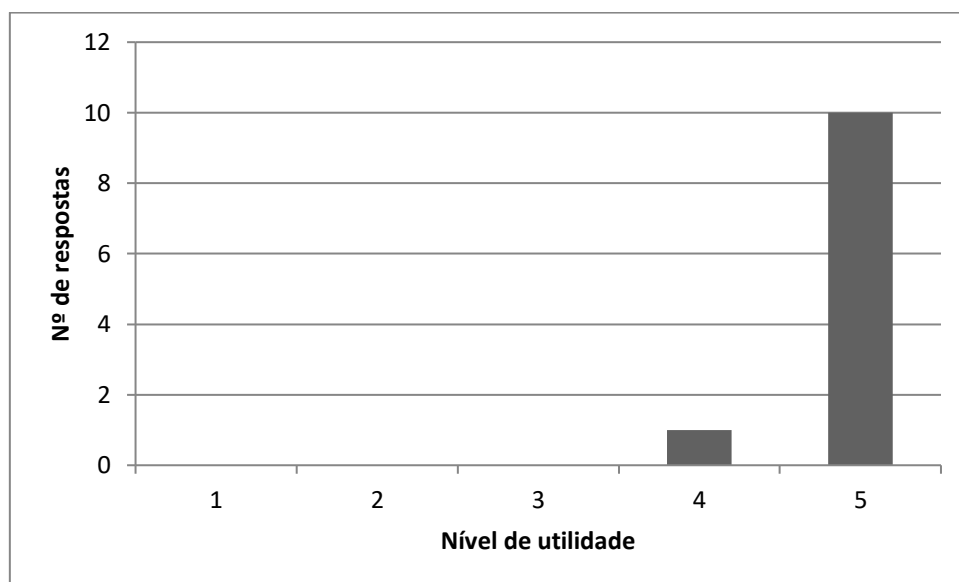


Figura 44. Nível de utilidade dos tutoriais para iniciantes

Quanto às funcionalidades existentes no protótipo (sugestões de produtos, ajuda no processo do *checkout*, personalização de pesquisa, respostas a eventuais dúvidas), 45,5% dos utilizadores (5) gostaram bastante, enquanto que outros 45,5% gostaram das funcionalidades (Figura 45).

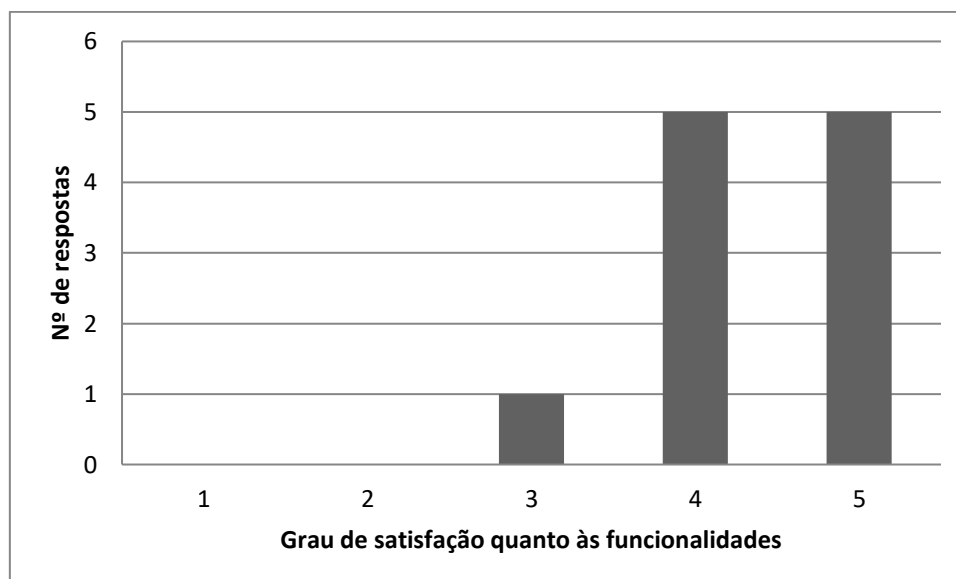


Figura 45. Número de respostas quanto ao grau de satisfação dos utilizadores

A maioria dos utilizadores – 72,8% (8 utilizadores) – concordam completamente que uma conversacional híbrida testada, é vantajosa em relação a uma interface unicamente gráfica, 18,2% (2 utilizadores) consideram-na vantajosa e apenas 1 utilizador (9%) é neutro quanto à sua opinião (Figura 46).

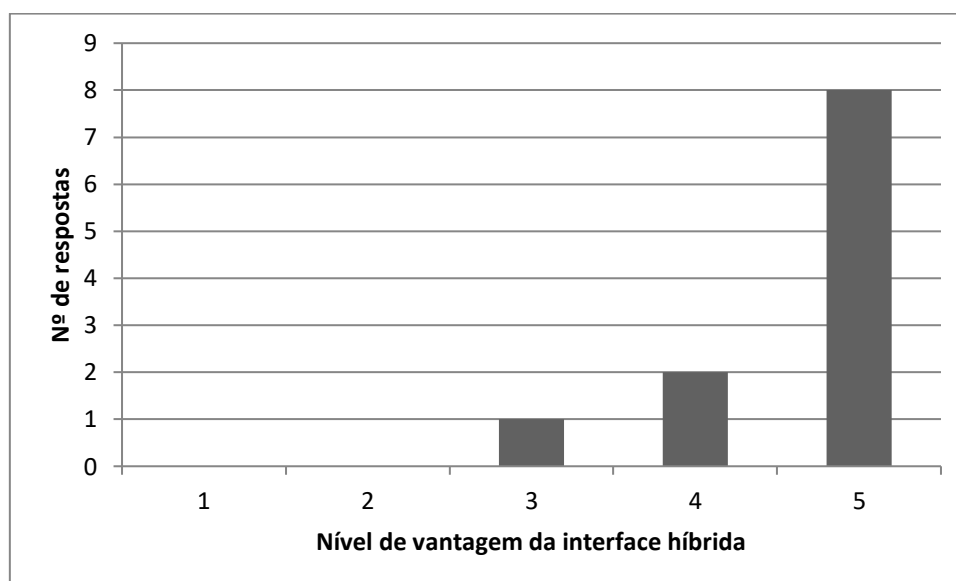


Figura 46. Nível de vantagem de uma interface híbrida em comparação a uma gráfica

A interação por voz também agrada os utilizadores e tendo como base as respostas destes, 54,6% (6 utilizadores) concordam completamente que a interação por voz é oportuna para certos contextos, 34,4% concordam que é útil e apenas 1 utilizador (9%) é neutro quanto à sua opinião (Figura 47).

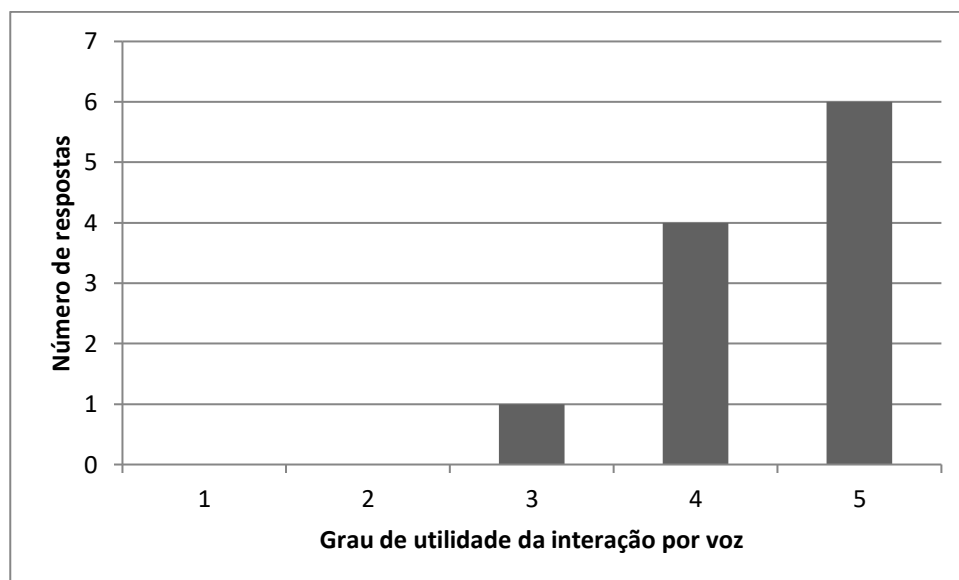


Figura 47. Grau de utilidade da interação por voz para diferentes contextos

Na segunda secção, foram definidas quatro questões qualitativas de resposta aberta, tal como é possível verificar na Tabela 11.

Tabela 11. Questões qualitativas da segunda secção

Questões
1. O que mais lhe agradou?
2. O que poderia ser melhorado na interface testada?
3. Alguma vez já realizou compras online?
4. Utilizaria esta interface híbrida ou invés da gráfica?

Uma das questões de resposta aberta tinha como finalidade obter um *feedback* do utilizador em relação ao que mais gostou na interface, obtendo diversas respostas, tal como é possível consultar na tabela abaixo.

Tabela 12. Respostas do que mais agradou ao utilizador na Shelf AI

ID Utilizador	O que mais agradou ao utilizador
#01	Ser uma interface híbrida, permitindo assim que escolher entre voz ou texto de acordo com a preferência e a forma como os fluxos estão organizados para as várias funcionalidades.
#02	Escolha de marcas.
#03	Combinação de funções: voz e escrita. Posso não querer dizer alguma coisa em voz alta.
#04	Rapidez nas respostas, diálogo personalizado e fluente.
#05	Possibilidade de pesquisar por comandos de voz e a naturalidade das respostas recebidas pois não são tão mecânicas como é habitual neste tipo de interações.
#06	Uso de linguagem natural.
#07	Rapidez da pesquisa.
#08	Quando guarda as nossas preferências e explica como devemos prosseguir.
#09	A disposição e composição da interface.
#10	Permitir comprar o que quero por voz sem ter de recorrer à escrita.
#11	A facilidade de pesquisa dos produtos que pretendia, utilizando os comandos de voz. A pesquisa é fácil e certa o que traz facilidade no momento da compra.

Com base nos resultados da Tabela 12, conclui-se que o que mais agradou à maioria dos utilizadores foi a possibilidade de interagir de forma híbrida, mas destacou-se a interação por voz devido à sua rapidez de resposta. Outros utilizadores gostaram essencialmente do

“atendimento” personalizado, pois o sistema guarda as preferências do utilizador e utiliza-as para apresentar de forma personalizada os resultados dos produtos.

Já na Tabela 13 estão retratadas as respostas dadas pelos utilizadores em relação à questão de resposta aberta, que tinha como intuito obter *feedback* quanto ao que pode ser melhorado na interface.

Ambas as tabelas estão a representar as opiniões dos utilizadores, que estão representados por um identificador. Isto permite uma comparação e relação do que determinados utilizadores gostaram mais e das melhorias que aconselham.

Tabela 13. Sugestões de melhorias dos utilizadores em relação à Shelf AI híbrida

ID utilizador	Sugestões de melhoria
#02	Poderia ter um contacto de ajuda ao cliente, caso um utilizador fique “encurralado” e não consiga continuar a interagir com a vertente áudio.
#03	Depois do produto seleccionado, a interface devia voltar a um “menu principal”.
#05	Não me agradou ser necessário dizer “Shelf” para ativar o modo de diálogo. O modo de diálogo podia permanecer ativo quando se faz a pesquisa e a navegação por voz.
#08	Penso que as funcionalidades que estão previstas são as ideais, só precisam de ser mais desenvolvidas.
#10	Deveria ser adaptado a qualquer língua.
#11	Poderia dizer o preço de outras marcas do mesmo produto que procuro.

É de referir que na tabela apenas estão expostas as respostas dos utilizadores que contribuíram com a sua opinião. As respostas levaram à conclusão de que a interface precisa de ser melhorada, nomeadamente naquilo que acham que deve ser acrescentado e não melhorado.

Por fim, quando foi questionado ao utilizador se utilizaria uma interface conversacional em vez de uma gráfica para realizar compras, todos os utilizadores mostraram preferência por esta interface. Alguns destacaram que utilizariam a interface conversacional híbrida devido às suas vantagens, enquanto que dois utilizadores mencionaram que escolhiam este tipo de interface dependendo da situação.

## 4.4. Síntese do Capítulo

Neste capítulo foi apresentado todo o processo de realização dos protótipos e como foram definidos os testes de usabilidade. Inclui também os resultados desses mesmos, assim como os questionários pós-teste.

Inicialmente foi realizada uma pesquisa quanto às possíveis ferramentas existentes para realizar protótipos de interfaces conversacionais híbridas, onde se conclui que não existe nenhuma ideal para este tipo de interfaces. Ferramentas como a Api.ai, Wit.ai e ChatScript são ideais para implementar o diálogo e numa primeira fase, sem recorrer à programação, enquanto que a SaySpring é uma ferramenta de prototipagem para interfaces de voz, onde é possível realizar testes sem programar. Apesar das suas boas funcionalidades e características, nenhuma foi selecionada para a realização dos protótipos, pois com estas não seria possível aplicar a componente gráfica da aplicação.

Deste modo, foi necessário recorrer a outras soluções, pelas quais se optou por utilizar o método de testes *Wizard of Oz*. Para realizar os protótipos funcionais através de clicks, foi utilizada a plataforma de prototipagem *online* Marvel, com o intuito de tratar da componente visual da aplicação assim como da interação manual. Para tratar da componente conversacional, foi decidido gravar previamente as *prompts* do sistema, realizadas no TTS da Oddcast e gravadas recorrendo ao *software* Audicity. Estas *prompts* foram organizadas estrategicamente de acordo com as tarefas definidas previamente, para durante o teste reproduzi-las de forma eficaz. Com o método WoZ, foi então possível simular a interação por voz, assim como o reconhecimento de voz. Através de cliques nos botões invisíveis criados nos protótipos, o responsável do teste tinha a possibilidade de manipular, através do TeamViewer, o que acontecia na interação, em todos os momentos onde seria necessária a inteligência artificial. Quanto às *prompts*, estas também foram reproduzidas pelo responsável, através do *software* SoundWire Server, realizando *stream* do computador para o *smartphone*.

As tarefas foram definidas ao mesmo tempo que os protótipos estavam a ser realizados, de forma a delimitar o que era necessário para estes. Os testes de usabilidade tinham como objetivo avaliar a eficácia da interface, a adequação das *prompts* e a utilidade da conversação por voz, assim como eventuais problemas de compreensão do utilizador e como se adaptava a uma interação híbrida. Em conjunto, foram também realizados questionários, de forma a se obter *feedback* dos utilizadores em relação a vários pontos da interface.

Concluiu-se então que a maioria dos utilizadores mostraram um elevado interesse em utilizar uma interface conversacional híbrida, ao invés de uma interface gráfica. Estes consideraram que as *prompts* da interface são naturais, que o tutorial é útil para iniciantes e mostraram uma elevada satisfação nas funcionalidades da aplicação, onde a que mais se destacou foi a pesquisa de produtos de forma rápida através da voz.

## 5. Conclusões e Trabalho Futuro

Esta dissertação tinha como meta responder aos objetivos e questões de investigação inicialmente definidos.

Após uma análise, realizada no Capítulo 2, referente às diferentes interfaces conversacionais existente,s assim como às interfaces gráficas associadas ao *e-commerce*, foi possível compreender como estas apresentam certas limitações para a realização de compras. De seguida, decorreu um estudo sobre o *design* para interfaces conversacionais a fim de compreender quais os pontos fundamentais para criar uma interação conversacional agradável para os utilizadores. Posto isto, foi possível ter conhecimento da importância dos fatores humanos, para não induzir a experiência de interação a algo que possa tornar-se frustrante para o utlizador. Existe esta preocupação em particular com os fatores humanos, pois é necessário ter em conta a capacidade de memória cognitiva, na fase do *design*. Esta capacidade pode interferir bastante na interação do utilizador, e por isso é necessário ter em consideração o seu limite, de forma a criar estrategicamente a componente de diálogo da interface. Essa limitação criou a necessidade de desenhar *prompts* curtas, pois existe o risco do utilizador apenas se lembrar das primeiras palavras da frase. Outro aspeto exterior aos fatores humanos, mas que também pode ser crucial na satisfação do utilizador está relacionado com o *feedback* do sistema. Esse *feedback* permite manter o utilizador sempre informado do que está a acontecer com a interface, reduzindo assim a possibilidade de existir receio e dúvida e potenciando a uma interação mais fluída. Tendo em conta estes aspetos e certas regras e princípios de *design*, foi possível adquirir certas bases para ter consciência de como é possível criar uma oportunidade de interação agradável nas interfaces conversacionais, principalmente na interação por voz, uma vez que é a menos comum.

Tendo em consideração a análise realizada previamente, quanto às interfaces conversacionais e gráficas de *e-commerce* existentes, assim como os aspetos mais relevantes no *design* conversacional, foi possível iniciar com bases suficientes, a fase de *design* da Shelf AI. Nesta fase, durante a criação do fluxo do diálogo, foi perceptível a forma como os utilizadores poderiam interagir com outras pessoas no contexto de compras, com a finalidade de auxiliar na

criação do fluxo e das *prompts* do sistema. Após a escolha de determinados produtos, foi compreensível que as necessidades dos utilizadores variam de acordo com os produtos e por isso o diálogo tem de ser diferente nas pesquisas de acordo com cada produto. Quanto a esta e outras funcionalidades, também se concluiu que a interação por voz não é vantajosa em todos os momentos de interação, recorrendo assim à componente gráfica para auxiliar a comunicação para com o utilizador.

A análise destas fases, a realização dos testes de usabilidade quanto ao protótipo desenhado, assim como os questionários realizados, permitiram em conjunto atingir os objetivos e adquirir respostas às questões de investigação.

Concluindo, as interfaces conversacionais híbridas são úteis, pois o utilizador pode interagir de acordo com as suas necessidades, pode ser adaptável para pessoas com limitações e permite pesquisas rápidas e eficazes. Devido à sua componente de inteligência artificial, também pode interagir de forma personalizada, tendo assim a capacidade de adaptar-se a diferentes contextos de utilização. O utilizador espera obter respostas com uma estrutura natural, no entanto, com os testes de usabilidade, foi notório que o próprio utilizador, ao interagir com a interface, apresenta ainda certa tendência em interagir de forma mais mecânica, com receio de que a interface possa não reconhecer conteúdo que não seja diretamente do contexto. O utilizador ao interagir com uma interface conversacional de *e-commerce*, espera obter uma pesquisa rápida, eficaz e personalizável, ou seja, que vá de encontro às suas necessidades. Espera igualmente que uma interface destas, possa retirar qualquer dúvida que tenha em relação a um produto, ou por exemplo ao estado da sua encomenda.

Com os testes de usabilidade chegou-se à conclusão de que o *design* conversacional criado para a Shelf AI satisfaz estes diferentes pontos apreciados pelos utilizadores. Foram poucas as indicações dos utilizadores para potenciais melhorias da interface e a maioria não está relacionada com o foco principal, a componente conversacional, mas sim com novos pontos a adicionar, mostrando assim que todas as análises e *design* realizado vão de encontro aos objetivos.

No futuro, existem novos objetivos definidos para tornar a componente conversacional mais completa e com potenciais novas funcionalidades, de forma a que a Shelf AI fique o máximo idealizada para ser implementada pela Xarevision quando possível.

Uma das alterações passa por alterar o comportamento da interface quando é adicionado um produto ao carrinho, que tal como sugerido por um dos utilizadores dos testes, ao adicionar um produto, a aplicação de seguida deve ser reencaminhada para uma espécie de página inicial, não ficando situada num ecrã que apresente o produto anteriormente adicionado.

Outra das alterações passa por preparar a aplicação para várias línguas, levando a uma reformulação das *prompts* a nível da língua e semântica.



Uma funcionalidade já idealizada mas que não foi abordada pela impossibilidade de testar, passa pelo utilizador poder requisitar receitas, assim como o sistema apresentar sugestões de receitas quando deteta um certo número de produtos adquiridos pelo utilizador. Juntos, os produtos podem originar uma determinada receita, que vai de acordo com o perfil do utilizador. As receitas poderão ser apresentadas visualmente e em simultâneo por voz, onde o utilizador poderá pedir apenas pelos ingredientes, por exemplo caso queira comprá-los, assim como pedir os passos da receita no momento em que precisa.

Tendo em conta o conhecimento e experiência obtida com a realização desta dissertação, um dos objetivos pessoais passa por continuar a trabalhar na área do *design* conversacional, caso o futuro permita, pois foi uma experiência nova e agradável. Como já mencionado, a área das interfaces conversacionais está em constante evolução, por isso, gostaria de contribuir para essa evolução trabalhando na área do *design*.

## 6. Bibliografia

- (21 de novembro de 2016). Obtido em 10 de Maio de 2017, de Instituto Nacional de Estatística:  
[https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine\\_indicadores&indOcorrCod=0002972&contexto=bd&selTab=tab2](https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0002972&contexto=bd&selTab=tab2)
- Abdul-Kader, S., & Woods, J. (2015). Survey on *Chatbot* Design Techniques in Speech Conversation Systems. *International Journal of Advanced Computer Science and Applications*.
- Allen, J., D. Byron, M. D., Ferguson, G., Galescu, L., & Stent, A. (2000). *An Architecture for a Generic Dialogue Shell*.
- Amazon Echo*. (s.d.). Obtido em 1 de Setembro de 2016, de Amazon:  
<https://www.amazon.com/Amazon-Echo-Bluetooth-Speaker-with-WiFi-Alexa/dp/B00X4WHP5E>
- Barker-Plummer, D. (26 de junho de 2012). *Turing Machines*. Obtido em 15 de dezembro de 2016, de Stanford Encyclopedia of Philosophy: <https://plato.stanford.edu/entries/turing-machine/#Describing>
- Barker-Plummer, D. (2016). Turing Machines. In *The Stanford Encyclopedia of Philosophy*.
- Baymard Institute. (s.d.). *37 Cart Abandonment Rate Statistics*. Obtido em 12 de dezembro de 2016, de <http://baymard.com/lists/cart-abandonment-rate>
- Baymard Institute. (s.d.). *E-Commerce Checkout Usability*. Obtido em 12 de dezembro de 2016, de <http://baymard.com/checkout-usability>
- Baymard Institute. (s.d.). *Top 50 E-Commerce Checkouts*. Obtido em 12 de dezembro de 2016, de <http://baymard.com/checkout-usability/benchmark/top-100>
- Beaudouin-Lafon, M., & Mackay, W. (2003). Prototyping Tools and Techniques. In *The Human-Computer Interaction Handbook*.
- Bickmore, T., & Cassell, J. (2000). *"How about this weather?" Social Dialogue with Embodied Conversational Agents*.
- Boves, L. (31 de Agosto de 2004). *Robust Conversational System Design*.

- Business Insider. (20 de setembro de 2016). *The Messaging App Report*. Obtido em 29 de setembro de 2016, de <http://www.businessinsider.com/the-messaging-app-report-2015-11?IR=T>
- Cassell, J., Bickmore, T., Campbell, L., Vilhjálmsdóttir, H., & Yan, H. (s.d.). *Conversation as a System Framework: Designing Embodied*.
- Chai, J., Horvath, V., Nicolov, N., Stys, M., Kambhatla, N., Zadrozny, W., et al. (2002). Natural Language Assistant: A Dialog System for Online Product Recommendation. *AI Magazine*.
- Chen, F., & Jokinen, K. (2010). *Speech Technology: Theory and Applications*. Springer.
- Cohen, M. H., Giangola, J. P., & Balogh, J. (2004). *Voice User Interface Design*.
- Colby, K., Hilf, F., Weber, S., & Kraemer, H. (1972). Turing-like Indistinguishability Tests for the Validation of a Computer Simulation of Paranoid Processes. *Artificial Intelligence*, 3, 199-222.
- Cooper, S. B. (2004). *The Incomputable Alan Turing*.
- Copeland, J. (maio de 2000). What is Artificial Intelligence?
- Cuello, J., & Vittone, J. (2013). Diseñando apps para móviles.
- Dix, A., Finlay, J., Abowd, G. D., & Beale, R. (2004). *Human-Computer Interaction* (3 ed.).
- Dohsaka, K., Asai, R., Higashinaka, R., Minami, Y., & Maeda, E. (s.d.). *Effects of Conversational Agents on Human Communication in Thought-Evoking Multi-Party Dialogues*.
- Dybkjaer, L., & Bernsen, N. O. (setembro de 2000). Usability Issues in Spoken Dialogue Systems. *Natural Language Engineering*, 6.
- Dybkjær, L., & Bernsen, N. O. (2001). *Usability Evaluation in Spoken Language Dialogue Systems*.
- Dybkjær, L., & Bernsen, N. O. (2001). *Usability Evaluation in Spoken Language Dialogue Systems*.
- Dybkjær, L., Bernsen, N. O., & Dybkjær, H. (1998). *A Methodology for Diagnostic Evaluation of Spoken Human-Machine Dialogue*.
- Fraser, N. (1997). Assessment of Interactive Systems. In D. Gibbon, R. Moore, & R. Winski, *Handbook on Standards and Resources for Spoken Language Systems* (3 ed., Vol. 24).
- Glass, J. (s.d.). *Challenges for Spoken Dialogue Systems*.
- Glass, J., Weinstein, E., Cyphers, S., & Polifroni, J. (2005). *A Framework for Developing Conversational User Interfaces*.

- Glass, J., Weinstein, E., Cyphers, S., & Polifroni, J. (2005). A Framework For Developing Conversational Users Interfaces.
- Gould, J. D., & Lewis, C. (1985). *Designing for Usability: Key Principles and What Designers Think* (Vol. 28). Communications of the ACM.
- Gould, J., & Lewis, C. (março de 1985). Designing for Usability: Key Principles and What Designers Think. 28.
- Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A., et al. (2004). *AutoTutor: A tutor with dialogue in natural language*.
- Güzeldere, G., & Franchi, S. (julho de 1995). Dialogues with colorful “personalities” of early AI. *Stanford Humanities Review archive*, 4(2), 161-169.
- Hern, A. (24 de março de 2016). *Microsoft scrambles to limit PR damage over abusive AI bot Tay*. Obtido em 16 de setembro de 2016, de <https://www.theguardian.com/technology/2016/mar/24/microsoft-scrambles-limit-pr-damage-over-abusive-ai-bot-tay>
- Hirschman, L., & Thompson, H. S. (1997). *Overview of Evaluation in Speech and Natural Language Processing*.
- Hofmann, H., Ehrlich, U., Reichel, S., & Berton, A. (s.d.). *Development of a Conversational Speech Interface Using Linguistic Grammars*.
- Isbister, K., & Doyle, P. (s.d.). *Design and Evaluation of Embodied Conversational Agents: A Proposed Taxonomy*.
- Kaplan, R. (2013). Obtido em 14 de dezembro de 2016, de <https://www.wired.com/2013/03/conversational-user-interface/>
- Kaplan, R. (2013). Obtido em 14 de dezembro de 2016, de <https://www.wired.com/2013/03/conversational-user-interface/>
- Klemmer, S. R., Sinha, A. K., Chen, J., Landay, J. A., Aboobaker, N., & Wang, A. (2000). *SUEDE: A Wizard of Oz Prototyping Tool for Speech User Interfaces*.
- Knight, W. (13 de junho de 2016). *Amazon Working on Making Alexa Recognize Your Emotions*. Obtido de <https://www.technologyreview.com/s/601654/amazon-working-on-making-alexa-recognize-your-emotions/>
- Knight, W. (13 de junho de 2016). *Amazon Working on Making Alexa Recognize Your Emotions*. Obtido em 15 de fevereiro de 2017, de MIT Technology Review: <https://www.technologyreview.com/s/601654/amazon-working-on-making-alexa-recognize-your-emotions/>
- Kopp, S., Gesellensetter, L., Krämer, N., & Wachsmuth, I. (s.d.). *A Conversational Agent as Museum Guide – Design and Evaluation of a Real-World Application*.

- Krug, S. (2000). *Don't Make Me Think*.
- Lee, C., & Lee, G. G. (s.d.). *Emotion Recognition for Affective User Interfaces using Natural Language Dialogs*.
- Lee, P. (25 de março de 2016). *Learning from Tay's introduction*. Obtido de <https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/#sm.0001yq2z734ocdkv11igkcz8qcu8s>
- Li, Q., Tur, G., Hakkani-Tur, D., Li, X., Paek, T., Gunawardana, A., et al. (s.d.). *Distributed Open-Domain Conversational Understanding Framework with Domain Independent Extractors*.
- Loisel, A., Duplessis, G., Chaignaud, N., Kotowicz, J.-P., & Pauchet, A. (14 de Março de 2014). *A conversational agent for information retrieval based on a study of human dialogues*.
- Lowdermilk, T. (2013). User-Centered Design.
- Lucente, M. (2000). Conversational interfaces for E-commerce applications. *Communications of the ACM*.
- Magic. (2015). *Personal Assistant On Demand*. Obtido de <https://getmagic.com/>
- Mctear, M. (s.d.). *Spoken Dialogue Technology: Enabling the Conversational User Interface*.
- Meng, H., Ching, P. C., Chan, S. F., Wong, Y. F., & Chan, a. C. (2004). *ISIS: An Adaptive, Trilingual Conversational System With Interleaving Interaction and Delegation Dialogs*.
- Messina, C. (19 de Janeiro de 2016). *2016 will be the year of conversational commerce*. Obtido em 13 de Setembro de 2016, de Medium: <https://medium.com/chris-messina/2016-will-be-the-year-of-conversational-commerce-1586e85e3991#.f02l6zaq7>
- Messina, C. (19 de janeiro de 2016). *2016 will be the year of conversational commerce*. Obtido de Medium: <https://medium.com/chris-messina/2016-will-be-the-year-of-conversational-commerce-1586e85e3991#.f02l6zaq7>
- Microsoft. (s.d.). *Cortana and privacy*. Obtido de <https://privacy.microsoft.com/en-us/windows-10-cortana-and-privacy>
- Milanesi, C. (6 de abril de 2016). *Alexa: A Short and Passionate Affair or A Long Standing Relationship?* Obtido em 8 de setembro de 2016, de Techpinions: <https://techpinions.com/alexa-a-short-and-passionate-affair-or-a-long-standing-relationship/44842>
- Milanesi, C. (22 de Junho de 2016). *Nobody puts Siri in a Corner: Why We might not See an Echo-like Device from Apple*. Obtido em 8 de Setembro de 2016, de Techpinions: <https://techpinions.com/nobody-puts-siri-in-a-corner-why-we-might-not-see-an-echo-like-device-from-apple/46375>

- Möller, S. (2005). Quality of Telephone-Based Spoken Dialogue Systems.
- Möller, S. (2005). *Quality of Telephone-Based Spoken Dialogue Systems*.
- Nass, C., Moon, Y., & Green, N. (1997). Are Machines Gender Neutral? Gender-Stereotypic Responses to Computers With Voices. *Journal of Applied Social Psychology*.
- Nass, C., Moon, Y., Fogg, B. J., Reeves, B., & Dryer, C. (1995). *Can computer personalities be human personalities?*
- Nielsen, J. (1 de janeiro de 1995). *10 Usability Heuristics for User Interface Design*. Obtido em 15 de fevereiro de 2017, de Nielsen Norman Group: <https://www.nngroup.com/articles/ten-usability-heuristics/>
- Nielsen, J. (4 de janeiro de 2012). *Usability 101: Introduction to Usability*. Obtido em 2 de dezembro de 2016, de Nielsen Norman Group: <http://www.nngroup.com/articles/usability-101-introduction-to-usability/>
- Nielsen, J., & Norman, D. (s.d.). *The Definition of User Experience*. Obtido em 2 de dezembro de 2016, de Nielsen Norman Group: <https://www.nngroup.com/articles/definition-user-experience/>
- Nielson, J. (1999). *Designing Web Usability: Designing for Simplicity*.
- Operator. (s.d.). Obtido em 15 de fevereiro de 2017, de Crunchbase: <https://www.crunchbase.com/organization/operator-2#/entity>
- Owda, M., Bandar, Z., & Crockett, K. (2007). *Conversation-Based Natural Language Interface to Relational Databases*.
- Pernice, K. (18 de dezembro de 2016). *UX Prototypes: Low Fidelity vs. High Fidelity*. Obtido em 15 de fevereiro de 2017, de Nielsen Norman Group: <https://www.nngroup.com/articles/ux-prototype-hi-lo-fidelity/>
- Pibernik, M. (abril de 2015). *An Analysis of Microsoft's 2015 'Siri vs Cortana' Advertising Campaign*.
- Rosset, S., Bennacef, S., & Lamel, L. (s.d.). *Design Strategies for Spoken Language Dialog Systems*.
- Rudnick, A. I., & Hauptmann, A. G. (1989). *Conversational interaction with speech systems*.
- Schnelle-Walka, D., & Lyardet, F. (2006). Voice User Interface Design Patterns. *Conference Paper*.
- Shapiro, A. (8 de agosto de 2016). Obtido em 14 de dezembro de 2016, de <http://www.hugeinc.com/ideas/perspective/conversational-interfaces-are-here>
- Shawar, B. A., & Atwell, E. (2007). *Chatbots: are they really useful?*

- Shneiderman, B. (1997). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Addison-Wesley .
- Tannenbaum, R. S. (1998). *Theoretical Foundations of Multimedia*.
- Traum, D., & Hinkelman, E. (dezembro de 1993). *Conversation Acts in Task-Oriented Spoken Dialogue*.
- Turing, A. (1950). Computing Machinery and Intelligence.
- Unger, R., & Chandler, C. (2012). A Project Guide to UX Design: For user experience designers in the field or in the making. New Riders.
- Wallace, R. S. (2009). The Anatomy of A.L.I.C.E. In *Parsing the Turing Test*.
- Wang, K. (2008). *Interface Adaptation for Conversational Services*.
- Weizenbaum, J. (1966). *Eliza: a computer program for the study of natural language communication between man and machine*.
- Yankelovich, N. (1996). *How Do Users Know What to Say?*
- Zanker, M., Gordea, S., Jannach, D., Jessenitschnig, M., & Bricman, M. (setembro de 2006). *Persuasive Online-Selling in Quality and Taste*.
- Zue, V. (2000). *Conversational Interfaces: Advances And Challenges*.
- Zue, V., & Glass, J. (agosto de 2000). *Conversational Interfaces: Advances and Challenges*.
- Zumbrunnen, A. (13 de Junho de 2016). *Technical and social challenges of conversational design*. Obtido em 24 de Agosto de 2016, de UX Design: <https://uxdesign.cc/my-website-is-now-conversational-here-is-what-i-learned-7e943cc6ace0#.trkp8c2ek>

# 7. Anexos

## Anexo A – Guião do Teste de Usabilidade

### Guião do Teste

Este teste de usabilidade tem como objetivo identificar problemas de usabilidade da aplicação, analisando como a interação entre o utilizador e a interface decorrem.

A interface a ser testada é uma interface híbrida onde pode **interagir** com esta através de **toque e voz**. A interação por voz permite um diálogo com linguagem natural para a ajudar a realizar compras de uma forma mais inata.

Nos momentos de interação em que não se encontre a decorrer um diálogo, para **ativar** o modo de **interação por voz**, da interface precisa pronunciar o comando “**Shelf.**”.

Durante o teste, necessita de ter em **atenção os produtos** que surgem nos **resultados**, para assim compreender quais as **possíveis escolhas para filtrar os resultados** consoante necessário.

Lembre-se que este teste é um **protótipo funcional sem recurso a programação**, existindo assim certos limites de interação.

#### Tarefa 1

Faça **login** na aplicação através do **facebook como se o seu nome fosse Diana ou Ricardo**. Quando surgir oportunidade pesquise por “**batatas fritas**”.

#### Tarefa 2

Pesquise por “**protetor solar**” e após concluir a filtragem, adicione ao carrinho o produto destacado.



### **Tarefa 3**

Pesquise por “**vinho**” e no fim da pesquisa, adicione ao carrinho o primeiro produto.

### **Tarefa 4**

Pesquise por “**iogurte sem lactose de morango**” e após concluir a pesquisa, adicione ao carrinho o produto de topo.

### **Tarefa 5**

Pesquise por “**água das pedras de limão**” e adicione ao carrinho o primeiro produto.

### **Tarefa 6**

Descubra qual é a **compota com menos calorias**.

### **Tarefa 7**

Realize o **checkout** e defina que a morada de entrega na **Rua das flores, número 1**. Marque a entrega para a **próxima terça feira pelas 15h00**. Caso não seja possível tente manter a **mesma hora para outro dia**.

## Anexo B – Inquérito Pós-Teste de Usabilidade

Classifique de 1 a 5 em que 1 é “não concordo” e 5 é “ concordo plenamente”, assinalando o número pretendido. Pode justificar a sua resposta sempre que quiser.

- 1. Considera que a linguagem do sistema, em relação à estrutura das frases, foi natural/humana?**

1	2	3	4	5
---	---	---	---	---

- 2. Considera útil existir uma espécie de tutorial integrado na primeira interação do utilizador com a interface? Tal como ocorreu no teste no momento da filtragem de produtos?**

1	2	3	4	5
---	---	---	---	---

- 3. Gostou das funcionalidades existentes (sugestões de produtos, ajuda no checkout, personalização de pesquisa, respostas a eventuais dúvidas)?**

1	2	3	4	5
---	---	---	---	---

- 4. Considera esta interface híbrida vantajosa em relação a uma interface unicamente gráfica?**

1	2	3	4	5
---	---	---	---	---

- 5. Considera a interação por voz oportuna para certos contextos e situações?**

1	2	3	4	5
---	---	---	---	---

Responda às seguintes questões:

**1. O que mais lhe agradou?**

---

---

---

**2. O que poderia ser melhorado na interface testada?**

---

---

---

**3. Alguma vez já realizou compras online?**

---

---

---

**4. Utilizaria esta interface híbrida ou invés da gráfica?**

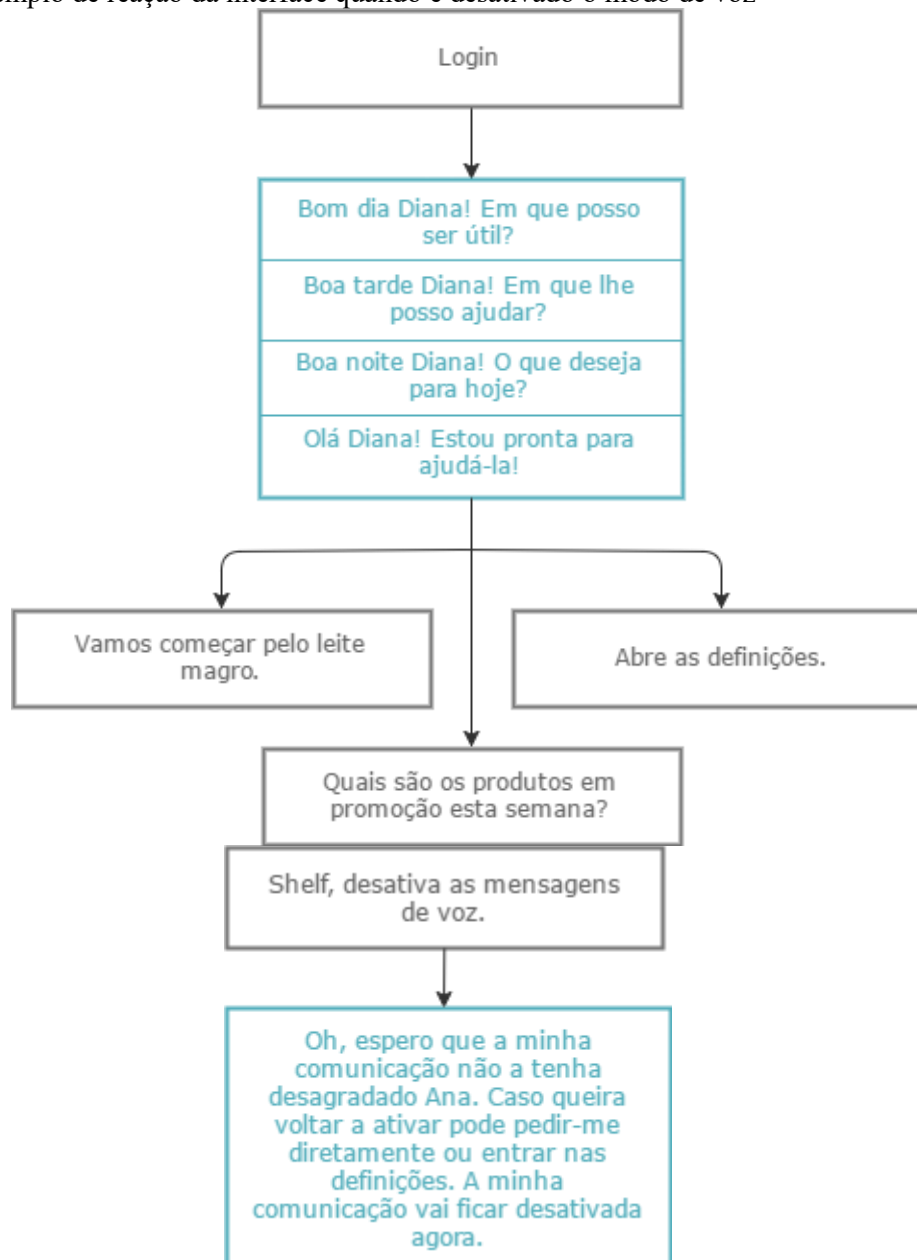
---

---

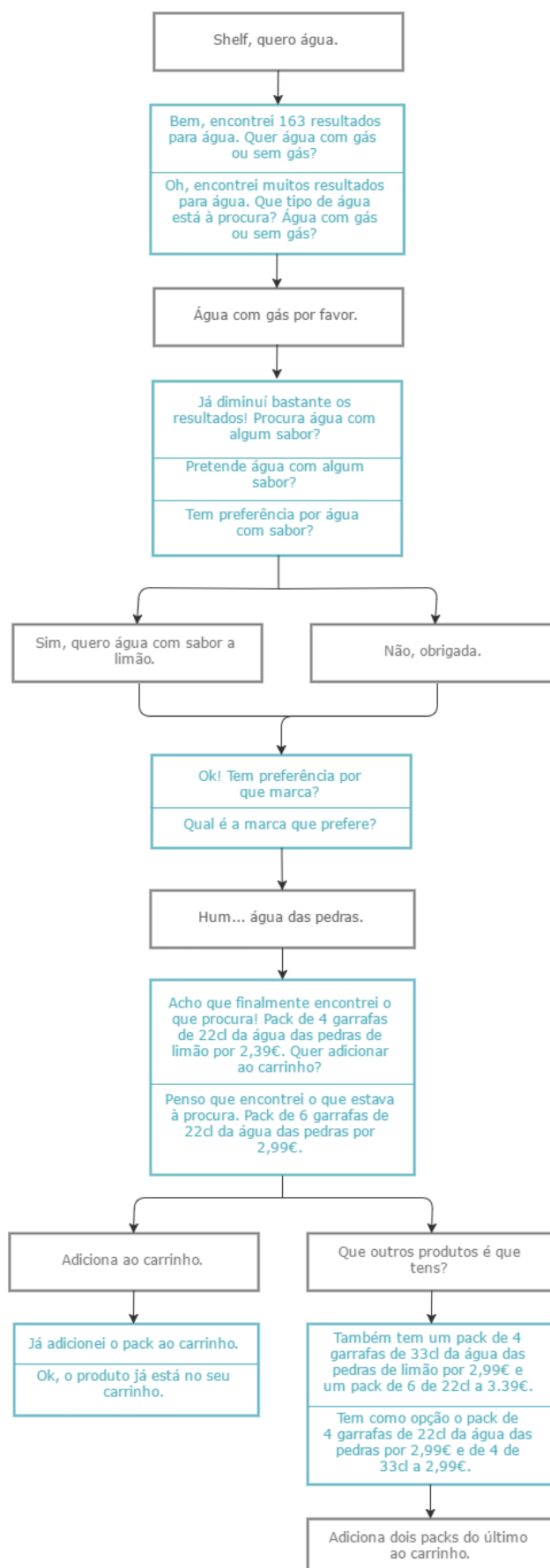
---

## Anexo C – Outros Fluxos de Diálogo

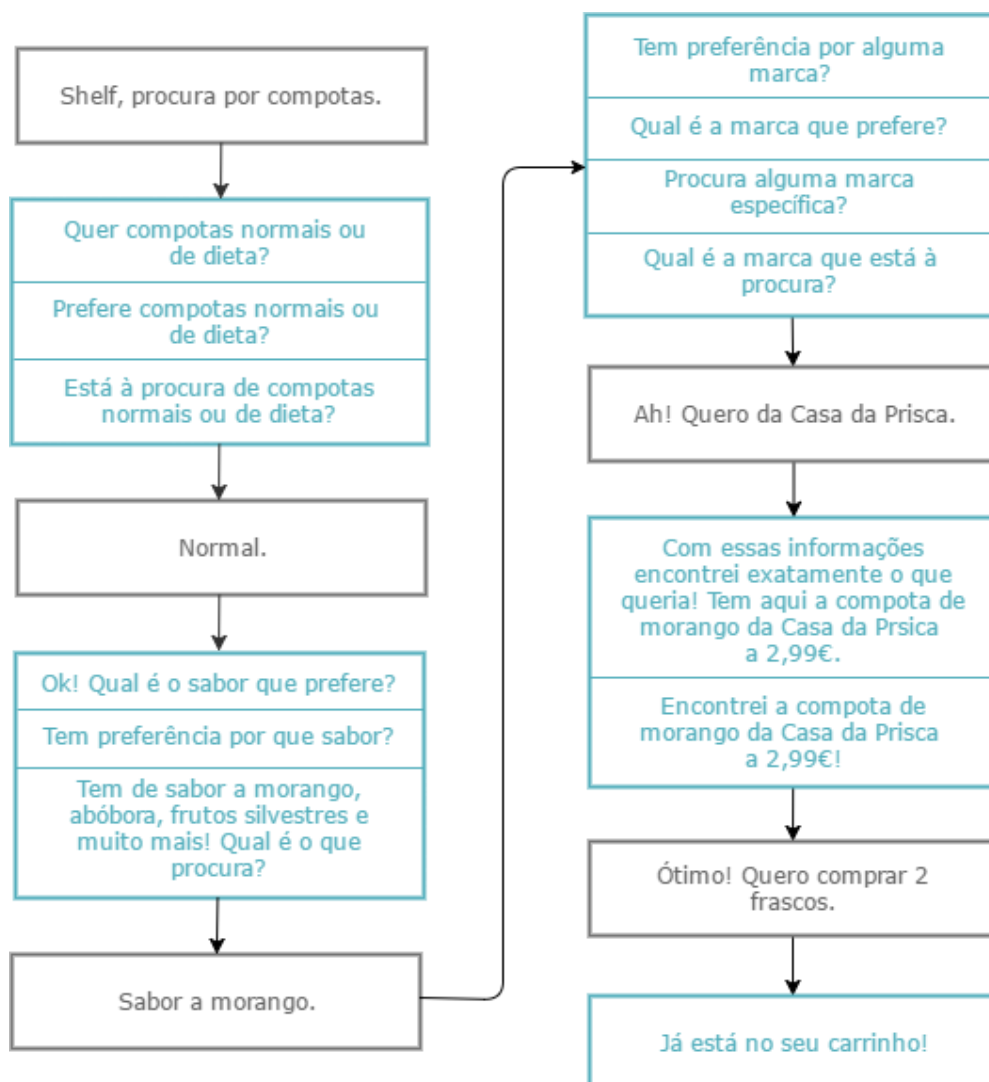
1- Exemplo de reação da interface quando é desativado o modo de voz



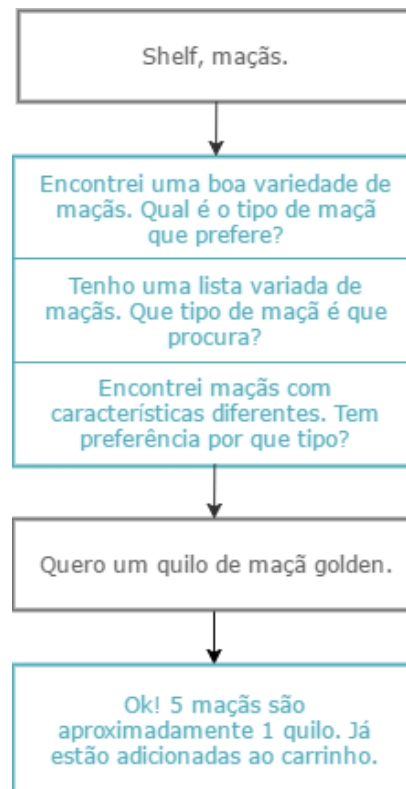
## 2- .Fluxo de como pode ocorrer todo o processo de pesquisa



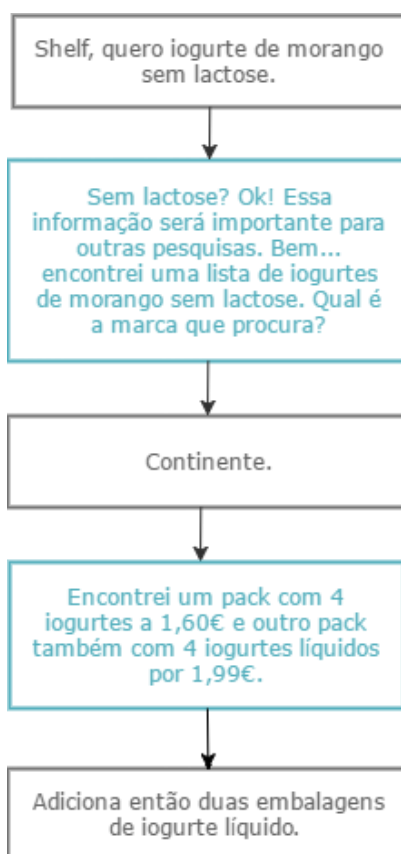
### 3 - Filtros definidos para a pesquisa de compotas



## 4 - Pesquisa de maçãs definida por apenas um filtro

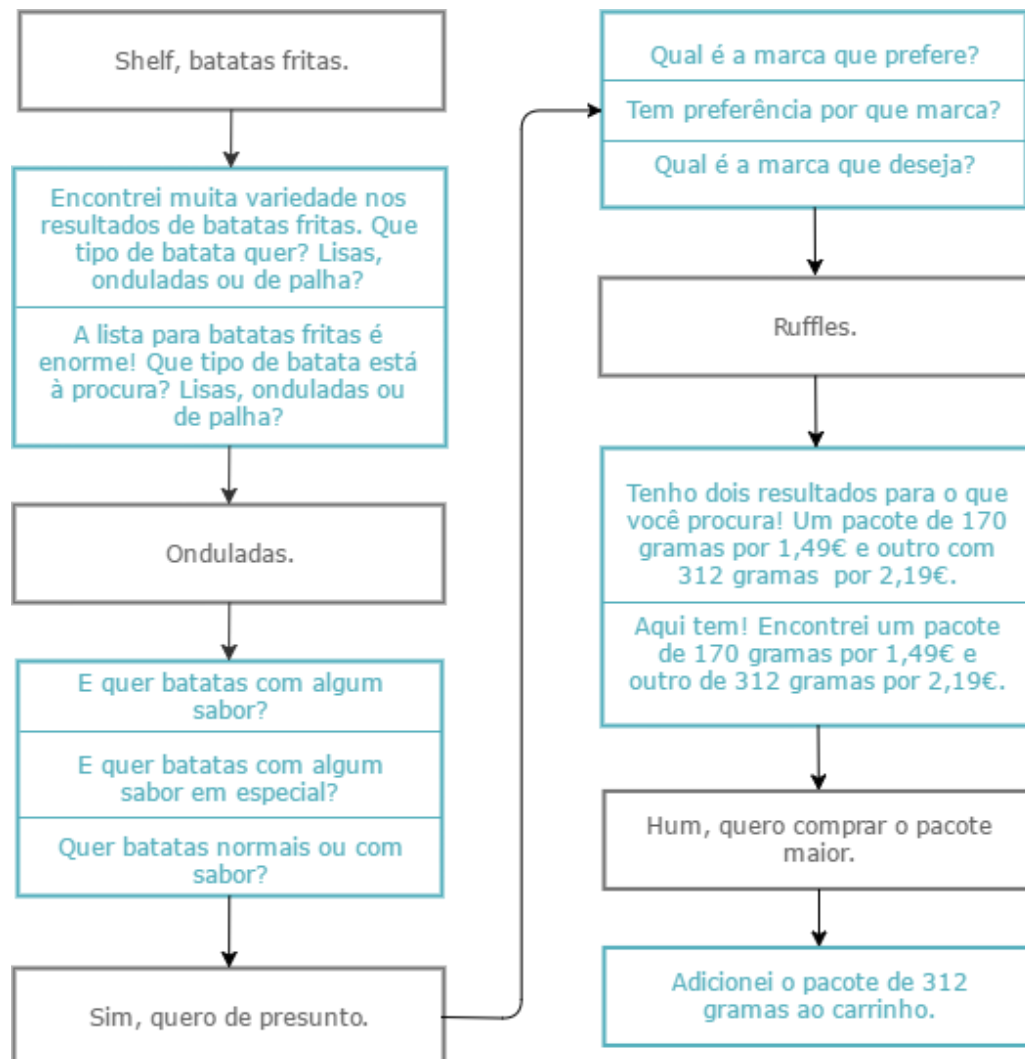


5 - Exemplo de pesquisa com entrada do produto quase completa

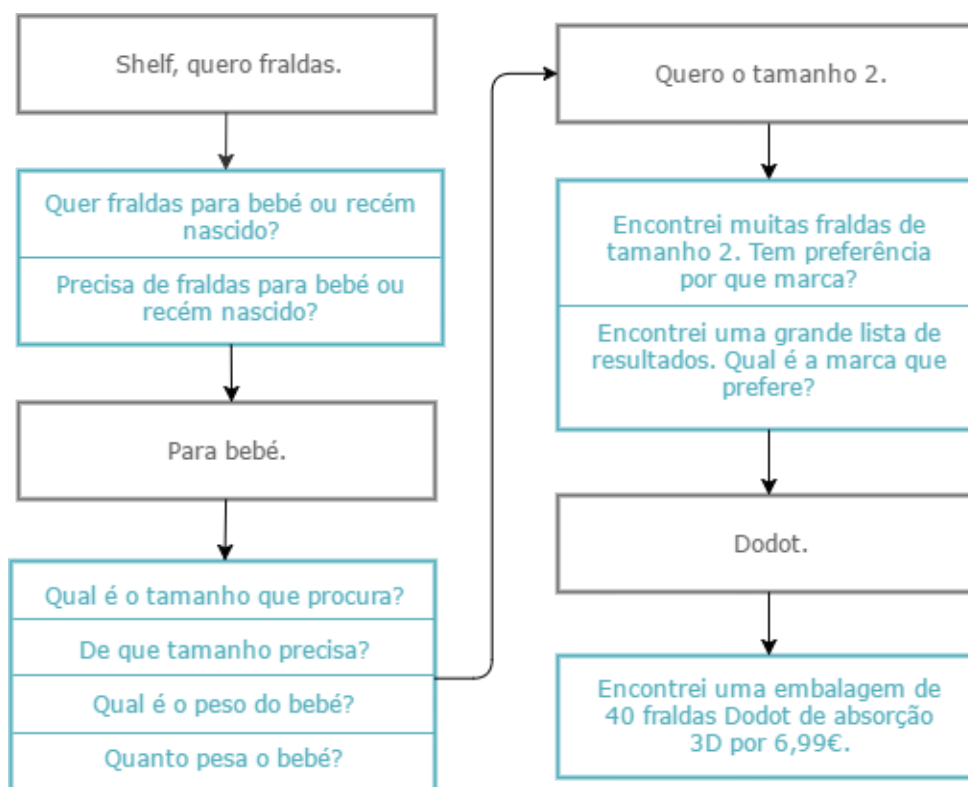




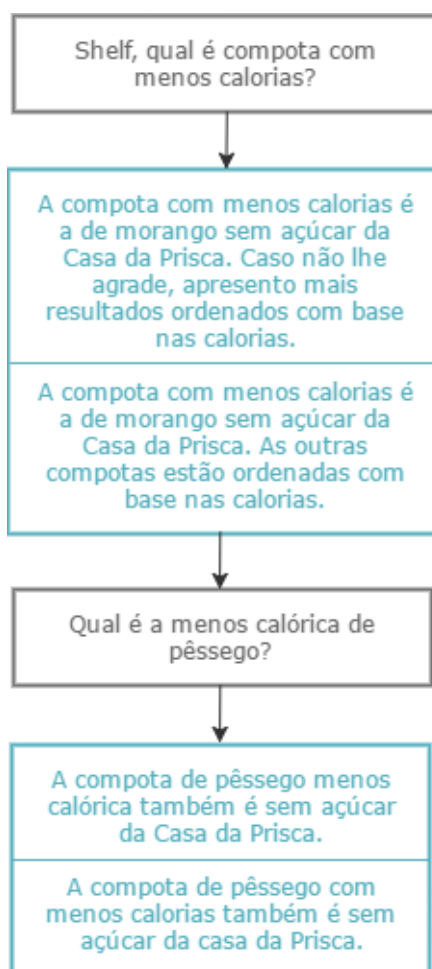
## 6 - Diálogo de pesquisa com utilização de todos os filtros para batatas fritas



7 - Pesquisa de fraldas com utilização de todos os filtros



## 8 - Exemplo de como a Shelf AI pode retirar dúvidas ao utilizador



9 - Exemplo de desvantagem ao mencionar por voz bastantes produtos

